Part 2

Lender-borrower relationships

Time:

March 2nd, Morning, 9:00～12:00;

March 2nd Afternoon, 14:00-16:00

Contents

Review of Finance (2009) 13: 261–307
doi: 10.1093/rof/rfn010
Advance Access publication: 24 April 2008

Banks, Distances and Firms' Financing Constraints

PIETRO ALESSANDRINI[1], ANDREA F. PRESBITERO[2] and
ALBERTO ZAZZARO[3]

[1]*Università Politecnica delle Marche and MoFiR;* [2]*Università Politecnica delle Marche, MoFiR and CeMaFiR;* [3]*Università Politecnica delle Marche, MoFiR and CFESPR*

Abstract. Bank deregulation and progress in information technology altered the geographical diffusion of banking structures and instruments, and reduced *operational distance* between banks and local economies. However, the consolidation of the banking industry promoted the geographical concentration of banking decision-making centres and increased *functional distance* between local banking systems and local borrowers. This paper focuses on the impact that these spatial diffusion-concentration phenomena had on the financing constraints of Italian firms over the period 1996–2003. Our findings show that greater functional distance stiffened financing constraints, especially for small firms, while smaller operational distance did not always enhance credit availability.

JEL Classification: G21, G34, R51

1. Introduction

Following the deregulation of credit markets and progress in information and communication technologies, geographical diffusion of banking structures and instruments has increased in almost every industrialized country. This trend has reduced *operational distance* between banks and borrowers and boosted competition in local credit markets. At the same time, due to the spectacular wave of bank mergers and acquisitions experienced in the banking industry, bank decision-making centres and strategic functions have been concentrated in few places within each

The paper benefited from comments made by participants at the conference on The Changing Geography of Banking (Ancona, September 2006), the XV International Tor Vergata Conference on Banking and Finance (Rome, December 2006), the 3rd Workshop on Economic Policy and Open Economy Macro (Milan, April 2007), the 43rd Conference on Bank Structure and Competition (Chicago, May 2007) and at seminars at the University of Bari, University of Naples Federico II, Università Politecnica delle Marche, University of Rome III and CFESPR (Portici). In particular, we wish to thank Arnoud Boot, Roberto de Bonis, Alfredo Del Monte, Valter Di Giacinto, Michele Fratianni, Adriano Giannola, Tullio Jappelli, Riccardo Lucchetti, Zeno Rotondi, Greg Udell, Francesco Venturini and two anonymous referees for helpful suggestions. We thank Antonio Riti (Capitalia) and Luigi Guiso for kindly providing us with Surveys on Italian Manufacturing Firms and data on blood donation in Italian provinces respectively.

country. The spatial concentration of banking power has greatly increased the *functional distance* of the control centre of lending decisions from local branches and borrowers.

Despite the ubiquity of these contrasting trends of spatial diffusion-concentration in the banking industry, their consequences for local economies have not yet been examined in depth. A large body of research has been devoted to studying the impact of the structure of local banking systems on local firms and economic development. However, these studies have failed to investigate the relative importance for lending decisions of distance within banks vis-à-vis distance between banks and borrowers. In this paper, we contribute to this literature by controlling for both operational and functional distances to provide a comprehensive account of the effects of the changing geography of local banking systems on local economies.

We focus on the effects of functional and operational distances on the financing constraints of Italian firms. We build aggregate indexes of operational and functional distance for local banking systems at the provincial level. Operational distance is measured with two proximity indexes (henceforth, operational proximity): the number of bank branches in the province in proportion to the local population or the provincial area. As a measure of functional distance we introduce a new index that takes account of the ownership structure of the local lending offices. Specifically, we calculate functional distance as the ratio of local branches weighted by, alternatively, the physical, economic and socio-cultural distance which separates them from the locus where their own bank is headquartered, to the total number of local branches. Moreover, we combine these three distance indicators by factorial analysis and obtain a composite latent measure of functional distance of the local banking system.

Four different econometric exercises are employed to quantify firms' financing constraints. The first two exercises are conducted at the firm level and evaluate the likelihood of credit rationing and the sensitivity of investment to cash flow. The third and fourth exercises are conducted at the market level and examine to what extent lines of credit granted to local firms are actually used as well as the flexibility of credit supply for local firms as proxied by the ratio of overdrawn to credit lines granted.

On the whole, our results corroborate the hypothesis that the lengthening of functional distance adversely affects credit availability to local firms; this effect is particularly significant for small businesses in the less developed provinces of southern Italy. In contrast, the effect of shortening operational proximity on firm financing constraints is ambiguous: in some specifications positive, while in others insignificant or outright negative. Interestingly, these results are consistent with those of Petersen and Rajan (2002), indicating decreasing importance over time of physical proximity between lenders and borrowers for the probability of loan approval.

A policy-oriented implication of our findings is that undifferentiated consolidation of the banking industry, not balanced by the presence of a strong local bank competitor, may be unable to drive competition towards the needs of small local firms, especially in peripheral regions. The negative externalities of increasing functional distance could be reduced by favouring a change in emphasis in bank organization from the search for economies of scale by standardized, arm's-length lending technologies, to economies of scope by making specialized credit instruments available to local firms (Berger and Udell, 2006).

While the theoretical reasons justifying the influence of functional and operational distances on credit allocation tend to be bank-firm specific, in this paper we adopt a market-based approach. Other studies have followed a market-based approach to investigate the effects of banking consolidation processes.[1] This approach has the advantage of being able to assess the impact of the changing structure of local credit markets on local borrowers, net of the reactions of incumbent banks. For example, the entry of banks in local credit markets may drive incumbents to focus more on their core lending market, where they enjoy informational advantages, so that financing constraints for local borrowers either increase or decrease according to the importance of that geographical market for incumbent banks (Hauswald and Marquez, 2003). Also, with increasing competition from outside, transactional lenders compel local lenders to leave a higher proportion of cash flow to their borrowers and thus may induce them to give up funding small but positive net present value projects (Inderst and Muller, 2007). However, a serious limitation of the market-based approach is that local firms could be actually indifferent to the structure of the local banking system. Insofar as firms are borrowing from bank branches located outside the province where they are headquartered, the correlation between our aggregate indexes of distance and firms' financing constraints would be spurious. Nevertheless, from Bank of Italy statistics, in Italy more than 90 per cent of credit to borrowers located in a province is granted by branches located in the same province. Thus, although we do not know the distribution of either subset, we can confidently assume that the phenomenon of out-of-market loans does not drastically affect our conclusions.

The paper proceeds as follows. Section 2 discusses the reasons why operational and functional distance should impact on lending relationships and selectively reviews the related empirical findings in the literature. Section 3 describes the data

[1] See, for example, Avery and Samolyk (2004), Collender and Shaffer (2003), Berger, Rosen and Udell (2007) for market-based studies on the US experience; Bonaccorsi and Gobbi (2001), Bonaccorsi and Dell'Ariccia (2004), Guiso et al. (2004a), Benfratello et al. (2008) for market-based studies on Italy.

and the distance variables. Sections 4, 5 and 6 display our four empirical exercises, illustrating for each of them the dependent and control variables, the empirical models and results. Section 7 discusses robustness checks and Section 8 concludes.

2. Why should operational and functional distances affect financing constraints?

2.1 OPERATIONAL PROXIMITY: THE THEORY

By operational proximity we refer to the physical proximity between the borrower and its lending office. The theoretical reasons why operational proximity may affect financing constraints for firms rest on (*i*) informational asymmetries underlying bank-firm relationships and (*ii*) transportation costs that fall on borrowers and shape competition in local credit markets.

Physical proximity to the local economy allows banks to supplement 'hard' data on borrowers with relevant 'soft' information collected locally. Soft information improves the quality of screening and monitoring borrowers, makes these actions less costly and facilitates relationship lending (Petersen and Rajan, 1994). Furthermore, informationally opaque borrowers may indirectly benefit from the presence of many competitors in local credit markets because local banks, in an effort to reduce the rigor of competition, redirect their lending towards borrowers for whom they have more informational rents (Boot and Thakor, 2000; Dell'Ariccia and Marquez, 2004).

However, close bank proximity to borrowers may also have negative consequences on financing constraints. Physical proximity translates into additional market power to the lending bank that can charge higher interest rates to close borrowers, either because of information rents (Hauswald and Marquez, 2006) or transportation costs (Degryse and Ongena, 2005). Furthermore, closer operational proximity may prompt "winner's curse" behaviour, inducing banks to be more conservative in terms of loan approval standards and interest rates (Broecker, 1990; Shaffer, 1998).

2.2 OPERATIONAL PROXIMITY: THE EVIDENCE

Empirical findings on the effects of physical proximity of banks to borrowers are mixed. At the bank level, Degryse and Ongena (2005) find that the physical proximity of borrowers to the lending office is associated with higher interest rates, as in spatial competition models. A negative relation between loan rates and distance is also found by Petersen and Rajan (2002) and Agarwal and Hauswald (2008) who, however, emphasise the role of proprietary information and firms' transparency. With regard to credit availability, Carling and Lundberg (2005) find no evidence of a relation between proximity and the probability that a bank may give a borrower a low rating. By contrast, Agarwal and Hauswald (2008) find that the closer a firm is to a bank branch the more likely that the loan application is approved. In the same vein, Brevoort and Hannan (2006) find that the probability of a bank lending in a given area decreases with physical distance from the nearest office of that bank.

At the market level, Avery and Samolyk (2004) report that the number of banks operating in a market is only weakly associated with small business lending growth in that market, whereas the number of offices has no impact at all on such a variable. Using Italian data, Bonaccorsi and Gobbi (2001) find that the ratio of branches to population in a province is positively associated with credit to firms located in that province, whereas it is negatively associated with bad loans. Benfratello et al. (2008) show that a higher branch density in Italian provinces increases the probability of firms introducing innovations, but reduces the sensitivity of investment to cash flow of small firms. However, Herrera and Minetti (2007), after controlling for the duration of the lending relationship, find that the effect of branch density on firm innovation is not significantly different from zero.

2.3 FUNCTIONAL DISTANCE: THE THEORY

By functional distance we refer to the distance between a local branch, where information is collected and lending relationships are established, and its headquarter, where lending policies and ultimate decisions are typically taken. From a theoretical point of view, the importance of functional distance for the lending policies of local branches has its roots in (*i*) the asymmetric distribution of information and the costs of communication within an organisation, and (*ii*) the economic, social and cultural differences across communities.

In recent years, it has become customary among banking scholars to emphasize the critical role of organizational issues for lending activities (Berger and Udell, 2002; Stein, 2002; Novaes and Zingales, 2004). Information on local borrowers, it is claimed, is in the hands of local bank managers and to a large extent this information is soft and not easily transferable to the upper echelons of the bank. This raises agency problems and stimulates behaviour that tends to hinder the collection and dissemination of soft information in large and hierarchically organised banks (Aghion and Tirole, 1997; Dessein, 2002; Harris and Raviv, 2005). Whatever the combination of delegation and centralisation of lending decisions chosen at the parent bank, significant organisational diseconomies arise that reduce the profitability of soft information loans (Stein, 2002; Takáts, 2004).

Others have emphasised effort and resources spent by a loan officer to influence decisions of senior managers at the headquarters of a large bank on the distribution of resources, job assignments and power within the organization (Milgrom and Roberts, 1990; Scharfstein and Stein, 2000). Such influence activities generate costs and credit misallocation for the bank. For example, to limit influence activities and induce accurate reporting on loan quality, loan officers are frequently rotated within the bank (Hertzenberg et al., 2007). The internal mobility of loan officers, in turn, reinforces incentives for the branch to produce short-term results. This may induce local managers to be too conservative about small and new borrowers who require a lot of soft information and to be too liberal about larger and well-established borrowers who are evaluated with hard information (Hirschleifer and Thakor, 1992; Palley, 1997; Berger and Udell, 2002).

The severity of agency costs and influence activities as well as their impact on credit allocation grow with the complexity of a bank's organisation. According to the literature, organisational diseconomies are shaped by bank characteristics such as size or organizational form (McAfee and McMillan, 1995; Berger et al., 1999; Berger et al., 2005). However, it is reasonable to assume that most organisational diseconomies are bank-firm specific and associated to the geographical dispersion of the bank organisation through branches and subsidiaries. In this respect, what matters is the distance between local and decision centres (Brickley et al., 2003; Mian, 2006), where distance encompasses physical, economic and socio-cultural attributes. For example, the greater the physical distance between the parent bank and local branch, the more costly it is to gather reliable information from loan officers, monitor their actions and enforce a lending policy designed at the bank's headquarters. At the same time, as differences increase in the economic structure of the regions where the parent bank and local branch are located, so does the informational rent of local loan officers who can accumulate specific knowledge of the local economy. Similarly, for branches located in socially and culturally diverse areas from bank headquarter, the bank-wide common business culture may succumb to the group-interaction effect as the root of individual behaviour.[2] In the absence of shared cultural values, senior bank officials and local managers find it hard to trust each other and communication between them becomes noisy, especially when dealing with soft information that is inaccessible to higher bank levels. In this vein, banks that open or acquire branches in a location where the culture and social environment drastically differ from those of its headquarters show a tendency to limit their activities to cherry-picking large and safe borrowers and selling asset management and other transactional products (Mian, 2006; Alessandrini et al., 2008).

[2] There is a huge literature on the importance of neighbourhood effects for disparate individual behaviours (from criminal actions to labour supply). An interesting application to banks is provided by Ichino and Maggi (2000) who show that shirking behaviours within a large Italian bank have significant regional differences due to group interactions and the different civicness which characterizes the area where bank branch employees work.

2.4 FUNCTIONAL DISTANCE: THE EVIDENCE

While there is robust evidence supporting the importance of agency and influence costs in banking organisations[3] and the disadvantages of large, complex bank organisations in small business and relationship-based loans[4], the role of functional distance on credit allocation has been much less extensively studied.

At the loan level, Mian (2006) shows that in Pakistan branches of foreign banks are less involved in small business and relational lending than branches of domestic banks, and that among the former the branches of non-Asian banks are the least involved. Jiménez et al. (2007) show that in Spain the likelihood of pledging collateral decreases with the distance between the lending branch and its bank headquarters, consistent with the hypothesis that functionally distant banks specialize in transaction-based loans. In the same vein, Casolaro and Mistrulli (2007) find that in Italy functional distant banks charge a lower interest rate to firms, especially when the latter are small sized. Finally, Liberti and Mian (2006) analyze a large multinational bank operating in Argentina and find that the sensitivity of the amount of credit facility granted to soft information available is statistically lower when the officer who approves the loan sits in a different location from that of the officer who collects the soft information.

At the bank level, Alessandrini et al. (2005) report that stand-alone banks headquartered in southern Italian regions are more involved in small business lending than southern banks affiliated to centre-northern holding groups. In a similar vein, Alessandrini et al. (2008) find that in Italy restructuring strategies pursued by the acquirer with respect to the asset portfolio of the acquired bank depend on the cultural distance between the provinces where the dealing partners are headquartered: the greater this distance, the stronger and more persistent is the decrease in acquired banks' loans to small business and the increase in transaction-based financial activities.

At the market level, Bonaccorsi and Dell'Ariccia (2004) and Collender and Shaffer (2003) construct a measure of what we call functional distance of the banking system from the local economy, taking into account the ownership structure of the local lending offices. Both papers find that the proportion of the local loan or deposit market held by in-market institutions (i.e., the functional proximity of the banking system) has some positive effect on local credit availability. With regard to the United States, Collender and Shaffer (2003) find that functional distance, under some circumstances, has a negative impact on local economic growth. Bonaccorsi and Dell'Ariccia (2004) find that in Italian provinces firm creation in industries

[3] Amongst others, see Udell (1989); Liberti (2004); Berger and DeYoung (2006); Liberti and Mian (2006); Hertzenberg et al. (2007).

[4] Amongst others, see Keeton (1995); Berger et al. (1998); Peek and Rosengren (1998); Sapienza (2002); Cole et al. (2004); Berger et al. (2005).

that are informationally opaque is positively associated with the share of deposits held by chartered banks in the same province.

In this paper, we make four contributions to this literature. First, we systematically address the issues of endogeneity of functional distance and omitted variables by using instrumental variables and system GMM estimations. Second, we focus on firms' financing constraints, whereas previous studies have analysed lending activities (Mian, 2006; Alessandrini et al., 2008), loan conditions (Casolaro and Mistrulli, 2007; Jiménez et al., 2007), firm birth rates (Bonaccorsi and Dell'Ariccia, 2004) and economic growth (Collender and Shaffer, 2003). Third, and more importantly, we introduce a new measure of local banking structure that takes into account the distance of local branches from their own headquarters. Previous studies have focused on the share of the local credit market controlled by out-of-market banks, based on the assumption that functional distance is a characteristic that banks either have or do not have, regardless of their location. Instead, we suggest that functional distance is better measured as a continuous property: all banks are subject to functional distance, albeit to different degrees. Although functional distance of a banking system is clearly correlated with the share of out-of-market owned branches, the two measures might lead to very different predictions. For instance, in our sample there are provinces where the share of out-of-market owned branches is the same, but the ultimate control of local branches is in the hands of banks that are headquartered at different distances from those provinces.[5] Finally, we control for other banking development and competition variables.

3. Data and variables

3.1 DATA

Our dataset relies on three sources. Firm-specific data come from a well-known survey on Italian manufacturing firms run every three years by the Observatory on SMEs by Capitalia, an Italian banking group (formerly Mediocredito Centrale).[6] This survey (henceforth, MCC survey) gathers information from a representative sample of firms with 11–500 employees[7] and from the universe of Italian manufacturing firms with more than 500 employees. Company balance sheet data are attached to the survey. We use the last three waves

[5] However, following the same argument, it is important to note that a bank-loan-level functional distance measure, which one cannot build with publicly available data, might lead to quantitatively different indications than our estimates based on aggregate functional distance measure.

[6] With regard to the banking literature, this dataset is employed, amongst others, by Benfratello et al. (2008), Herrera and Minetti (2007) and Huynh and Rotondi (2007).

[7] Firms are randomly selected and stratified by size (five classes of employees), industry (Pavitt classification) and geographical area (North and Centre-South).

covering the period 1995–2003. Taken together, they have 13,004 observations (9,674 with balance sheet data), but only 526 firms are present in all three waves due to the resampling rule followed in the survey. Data on bank and branch location, bank group composition, credit lines granted, credit used and overdrawn at the provincial level are from the Bank of Italy. Credit line information covers the period 1997–2003. Institutional and macroeconomic data at the provincial level on the functioning of the judicial system, value added and population are from the Italian National Institute of Statistics (ISTAT).

3.2 MEASURING DISTANCES

3.2.1 *Operational Proximity*

We use two measures of operational proximity of the banking system to a province j. First, we build a branch density index in terms of population, $OPERATIONAL\_PROXIMITY\_POPULATION = Branches_j \times 10{,}000/Population_j$. This is the most widely used index of bank presence in an area and allows straight comparison of our results with the literature. However, as has been noted by a referee, *OPERATIONAL_PROXIMITY_POPULATION* is only loosely associated with transportation and information costs embedded in our notion of operational proximity. A branch density index measured in proportion to the provincial surface area, $OPERATIONAL\_PROXIMITY\_AREA = Branches_j/Surface\ Area_j$, is an alternative and more straightforward proxy of transportation and information costs borne by borrowers and lenders.[8]

3.2.2 *Functional Distance*

Functional distance depends on ownership and organizational structure of banks operating locally. We compute functional distance of the banking system in province j by weighting the number of branches operating in that province by a distance indicator that captures the severity of informational and organizational frictions between the local branch and the headquarters of their parent bank:

$$FUNCTIONAL\_DISTANCE_j = \frac{\sum_{b=1}^{B_j} \left[Branches_b \times \ln\left(1 + D_{jz_b}\right)\right]}{\sum_{b=1}^{B_j} Branches_b},$$
$$\text{with } j, z = 1, \ldots, 103$$

[8] For example, in a standard circular city competition model with n branches equidistantly distributed over a circle of circumference S and linear transportation costs, in equilibrium a firm located at distance x from a branch i incurs transportation costs $t(S/n - x)$ to go to i or, equivalently, $t(1/OPERATIONAL\_PROXIMITY\_AREA - x)$, where t is the transportation cost per unit of length.

where B_j is the number of banks operating in the province j, $Branches_b$ is the number of branches belonging to the bank b, and D_{jz_b} is the distance indicator between the branch and its bank.

We build three alternative functional distance indexes: (i), $FUNCTIONAL\_DISTANCE\_KM_j$, where the weighting distance indicator D_{jz_b} is measured in physical distance between branch of bank b located in province j and its headquarter located in province z (KM_{jz_b}, with $KM_{jj_b} = 0$;) (ii) $FUNCTIONAL\_DISTANCE\_SOCIAL\_CAPITAL_j$, where D_{jz_b} is the absolute difference between the social capital of provinces j and z, $|SOCIAL\_CAPITAL_j - SOCIAL\_CAPITAL_{z_b}|$; and ($iii$), $FUNCTIONAL\_DISTANCE\_ECONOMIC\_STRUCTURE_j$, where the distance indicator is an index of economic dissimilarity between provinces j and z measured by shares of workers employed by m economic sectors, $\sum_{h=1}^{m} |W_{hj} - W_{hz_b}|$ for the absolute value.

Kilometric distances across 103 Italian provinces are calculated with respect to provincial capitals using an extension of Jenness' (2005) ArcView GIS software, Distance Matrix. Social capital is computed as the average voter turnout at the 21 referenda held in Italy in 1993, 1995 and 2001 as published by the Home Department.[9] The share of workers employed by sector h in province j is drawn from the VIII Industry and Services Census (ISTAT 2001). We consider 14 economic sectors (NACE one digit), but we checked the robustness of our results by using different levels of data desegregations and also reclassifying the share of manufacturing sectors according to the PAVITT classification.

Functional distance indexes are positively and highly correlated with one another and negatively correlated with operational proximity (Table I). This corroborates the idea that functional distance includes physical, cultural and economic distances; to take account of the simultaneous impact of these distances on functional distance, we combine *FUNCTIONAL_DISTANCE_KM*, *FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL* and *FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE* indexes into a single latent distance index (*FUNCTIONAL_DISTANCE_LATENT*) by taking the first component from a FActorial Matrices Analysis (FAMA; Fachin et al., 2002). This first component explains 81 per cent of the global variance of the three distance indexes.[10]

The underlying assumption of these functional distance indexes is that local branches are controlled by the chartered banks owning them, regardless of their

[9] See Guiso et al. (2004b) and Alessandrini et al. (2008). Following Guiso et al. (2004b), we also use the number of blood bags per million population voluntarily donated in each province to the Italian Association of Blood Donors (AVIS) in 1995.

[10] Only the first eigenvalue is greater than one, with the corresponding eigenvectors equal to 0.60, 0.58 and 0.55 for, respectively, *FUNCTIONAL_DISTANCE_KM*, *FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL* and *FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE*, and showing a similar contribution of the three distance indicators in the latent index *FUNCTIONAL_DISTANCE_LATENT*.

Table I. Banking distance variables, pairwise correlations

The table reports the pairwise correlations for the banking distance variables. * significant at 5%. Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j*. The weighting distance between the two provinces is measured: in kilometric distance (*FUNCTIONAL_DISTANCE_KM*), as the absolute difference in social capital (*FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL*), and as a dissimilarity index of the economic structure in terms of shares of workers employed by *m* economic sectors (*FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE*). A fourth index of functional distance (*FUNCTIONAL_DISTANCE_LATENT*) is obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The *FUNCTIONAL_DISTANCE* indexes are reckoned assuming that control of local branches is in the hands of the chartered banks that own them, regardless of their affiliation to a multibank holding company. *FUNCTIONAL_DISTANCE_KM_G*, *FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL_G*, *FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE_G* and *FUNCTIONAL_DISTANCE_LATENT_G* are built assuming that the holding company is in command of local branches. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches (× 10,000) over the surface area (*OPERATIONAL_PROXIMITY_AREA*)

	FUNCTIONAL_ DISTANCE_ KM	FUNCTIONAL_ DISTANCE_ SOCIAL_ CAPITAL	FUNCTIONAL_ DISTANCE_ ECONOMIC_ STRUCTURE	FUNCTIONAL_ DISTANCE_ LATENT	FUNCTIONAL_ DISTANCE_ KM_G	FUNCTIONAL_ DISTANCE_ SOCIAL_ CAPITAL_G	FUNCTIONAL_ DISTANCE_ ECONOMIC_ STRUCTURE_G	FUNCTIONAL_ DISTANCE_ LATENT_G	OPERATIONAL_ PROXIMITY_ POPULATION	OPERATIONAL_ PROXIMITY_ AREA
FUNCTIONAL_ DISTANCE_KM	1									
FUNCTIONAL_ DISTANCE_ SOCIAL_CAPITAL	0.8227*	1								
FUNCTIONAL_DISTANCE_ ECONOMIC_ STRUCTURE	0.6949*	0.6538*	1							
FUNCTIONAL_ DISTANCE_LATENT	0.9811*	0.9140*	0.7308*	1						
FUNCTIONAL_ DISTANCE_KM_G	0.8158*	0.7062*	0.5342*	0.8083*	1					
FUNCTIONAL_ DISTANCE_ SOCIAL_CAPITAL_G	0.7229*	0.8738*	0.4974*	0.7990*	0.8505*	1				
FUNCTIONAL_DISTANCE_ ECONOMIC_ STRUCTURE_G	0.6278*	0.5868*	0.8035*	0.6553*	0.7807*	0.6787*	1			
FUNCTIONAL_ DISTANCE_LATENT_G	0.7997*	0.7838*	0.5491*	0.8237*	0.9637*	0.9199*	0.7766*	1		
OPERATIONAL_ PROXIMITY_ POPULATION	−0.5618*	−0.6117*	−0.3872*	−0.5975*	−0.5591*	−0.6559*	−0.4281*	−0.6043*	1	
OPERATIONAL_ PROXIMITY_AREA	−0.1610*	−0.2335*	−0.1040*	−0.1917*	−0.1876*	−0.2650*	−0.1605*	−0.2194*	0.1609*	1

affiliation to a multibank holding company. As parent banks typically preserve a certain influence on the decisional process of affiliated banks, this is a conservative assumption that probably undervalues the impact of functional distance on financing constraints to firms.[11] Nevertheless, we will check the robustness of our findings by analyzing the opposite situation, in which the holding company is assumed to be in command of local branches, and we calculate functional distance indexes consistently (hereafter, *FUNCTIONAL_DISTANCE_G*). Moreover, since our results are qualitatively robust to changes in functional distance indexes, to save space in the sections below we primarily discuss regressions with *FUNCTIONAL_DISTANCE_KM* and *FUNCTIONAL_DISTANCE_LATENT*.

3.2.3 Descriptive Statistics

All distance measures show a great variability over time and across provinces (Table II). The increase in operational proximity and functional distance indexes from 1996 to 2003 was statistically significant and more so in southern provinces, where the geography of banking has had deeper changes than in the North of the country. The provincial distribution of distance indicators mirrors the well known Italian geographical divide between northern-central and southern regions: banks in the North and the Centre are functionally and operationally less distant than banks in the South. The North-South divide is even sharper when functional distance is computed with respect to the headquarters of parent banks.

3.3 MEASURING FINANCING CONSTRAINTS

Due to the lack of an established method in measuring financing constraints to firms, in this paper we have adopted a robust approach. We run different econometric exercises using four alternative indicators of financing constraints as dependent variable. The first two indicators are computed at the firm level from survey and balance-sheet data; the third and fourth at the market level. In the first exercise, we use self-reported answers to a question on credit rationing in the MCC Survey: a dummy variable, *RAT*, takes the value of 1 if the firm answered yes to the question

[11] In Italy the headquarters of parent banks are much more geographically concentrated than those of chartered banks. According to the Register of Banks and Banking Groups held by the Bank of Italy, by 2004, 148 chartered banks had their headquarters in one of the southern regions, whereas no banking group had their holding company or parent bank in the these regions. Among the 148 southern banks, 111 were small Credit Cooperatives, 21 were affiliated to Northern-Central banking groups and only 16 were independent, stand-alone banks. Considering only the distance within chartered banks, therefore, reduces the geographical and time variability of functional distance indexes and underestimates the distance to southern provinces where few holding groups are located (see Table II)

Table II. Banking distance variables: Summary statistics

The table reports the summary statistics for the banking distance variables. Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j* (*FUNCTIONAL_DISTANCE_KM*) and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches (× 10,000) over surface area (*OPERATIONAL_PROXIMITY_AREA*). For *FUNCTIONAL_DISTANCE_KM* we also report, in parenthesis, the mean and the standard deviation of the index built by weighting the number of branches by the actual kilometric distance between the province of the branch and that where the parent bank is headquartered. The last four rows report a two-tailed test of the null hypothesis that (1) the values of the banking distance variables in 1996 and in 2003 are equal, and (2) the averages of the banking distance variables over the period 1996–2003 in the North and in the Centre-South are equal. Actual differences are reckoned, with *** indicating a statistical significance at 1% level, while the t-statistics are reported in parenthesis.

Year	*FUNCTIONAL_ DISTANCE_KM*		*FUNCTIONAL_ DISTANCE_ LATENT*		*OPERATIONAL_ PROXIMITY_ POPULATION*		*OPERATIONAL_ PROXIMITY_ AREA*	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
1996	2.972 (133.14)	1.082 (94.97)	2.308	0.955	4.509	1.612	10.441	12.503
1997	3.035 (138.25)	1.094 (100.54)	2.363	0.969	4.656	1.677	10.839	13.023
1998	3.078 (141.92)	1.101 (102.02)	2.401	0.982	4.885	1.678	11.234	13.417
1999	3.158 (150.52)	1.116 (115.41)	2.474	1.003	5.087	1.734	11.621	13.855
2000	3.213 (154.67)	1.121 (122.52)	2.525	1.014	5.267	1.789	12.070	14.330
2001	3.277 (161.16)	1.137 (129.23)	2.581	1.033	5.449	1.843	12.516	14.832
2002	3.394 (172.53)	1.132 (135.06)	2.672	1.031	5.533	1.875	12.803	15.154
2003	3.451 (172.82)	1.151 (135.51)	2.720	1.038	5.538	1.904	12.957	15.301
1996–2003 Growth rates								
Italy	16.10		17.83		22.09		24.10	
North	15.31		16.40		21.95		25.77	
Centre	14.26		14.37		21.44		25.60	
South	17.57		20.37		26.46		17.48	
t-test on the differences between:								
1996 vs 2003	−0.478*** (9.755)		−0.411*** (9.426)		−1.029*** (15.901)		−2.516*** (8.287)	
North vs Centre-South	−0.903*** (4.544)		−0.899*** (5.150)		2.485*** (10.141)		8.808*** (3.316)	

"*In [the last year of the survey] would the firm have desired more credit at the interest rate agreed with the bank?*" and 0 otherwise (Capitalia, 2005).[12]

The second exercise tries to identify the effect of distances on financing constraints, following the large literature on the sensitivity of firm investment to cash flow (Fazzari et al., 1988).

In the third and fourth exercises, financing constraints are computed at the provincial level as: (*i*) the ratio of credit lines used by in-province borrowers to credit lines made available to them (*CREDIT_DRAWN*), and (*ii*) the ratio of overdrawn to credit lines available to in-province borrowers (*OVERDRAWN*). They are widely employed indicators of credit market tightness.[13] In many countries, including Italy, firms typically negotiate with banks credit lines that ensure a certain degree of flexible use. For example, borrowers have the option of either not fully utilizing the credit line, free of any charge, or overdraw up to a certain amount, subject to fees and penalty interest rates on the overdrawn. When the credit market is tight, firms have a lower credit buffer and use a higher proportion of the negotiated credit line (i.e., *CREDIT_DRAWN* is higher); moreover, the right to overdraw is narrower and the flexibility of credit supply smaller (i.e., *OVERDRAWN* is lower).

4. Credit rationing

4.1 THE ECONOMETRIC MODEL

We estimate the probability of a firm being credit rationed as a function of operational proximity, functional distance and other firm-specific and geographical control variables:

$$\begin{aligned}\Pr(RAT_{ij}) = f(&OPERATONAL\_PROXIMITY_j,\ FUNCTIONAL\_DISTANCE_j,\\ &FIRM_i, PROV_j)\end{aligned} \quad (1)$$

where i is the firm, *FIRM* stands for m firm-specific control variables and *PROV* for n control variables at the provincial level, which should affect the likelihood of being financially constrained.

4.1.1 Control Variables

All firm-specific variables come from the MCC survey and balance-sheet data. In our basic specification, *FIRM* includes variables such as firm size (*SIZE*),

[12] The survey contains two other questions concerning the firms' willingness to pay a slightly higher interest rate than the current one and the actual rejection of a loan application. Since the order and method in which these questions are reported change with the survey waves, responses are not perfectly comparable across the three datasets and we have decided to concentrate our analysis only on the more general question on credit rationing, though it mixes quantity with price rationing.

[13] Amongst others, credit drawn by borrowers is employed at the market level by Bonaccorsi (2003) and at the firm level by Agarwal et al. (2006) and Bonaccorsi and Gobbi (2007).

propensity to innovate (*R&D*), return on investment (*ROI*), degree of indebtedness (*DEBT*) and bank-firm relationships (*BANK_NUMBER* and *BANK_PROVINCE*).

SIZE is measured by the logarithms of the current number of employees and is expected to be negatively associated with the likelihood of credit rationing. The profitability of firms, *ROI*, is computed as gross operating earnings on invested capital and is expected to reduce the probability of being rationed. *DEBT* is measured by the logarithm of the debt to equity ratio. As usual, we assume that it is positively correlated with firm riskiness and hence expect that the higher *DEBT* the higher the probability of being credit rationed. *R&D* is computed as the share of workers employed in Research and Development departments on the total labour force. Following the literature, we expect that more innovative firms are likely to be credit rationed, because they are more informationally opaque and perceived as riskier by banks. To mitigate possible endogeneity problems all these variables are pre-dated to the starting year of the survey.

Finally, among the firm-specific controls we include two indicators of the bank-firm relationship: (*i*) the number of banks with which the firm does business (*BANK_NUMBER*) and (*ii*) a dummy variable which takes a value of 1 if the firm's main bank is headquartered in the same province and 0 otherwise (*BANK_PROVINCE*). The *a priori* expectations on both *BANK_NUMBER* and *BANK_PROVINCE* are ambiguous. The former is a measure of multiple lending. This, in one way, should reduce financing constraints by putting lending banks in competition with each other but, in another way, could exacerbate financing constraints for firms by lessening banks' incentives to engage in relational lending. In the same way, relying on a physically close bank could help to overcome information asymmetries. But it could also make the firm more likely to be informationally captured by the local bank with negative effects on its financing constraints (Ongena and Smith, 2000).

PROV includes control variables for both the characteristics of the banking system and the institutional environment at the provincial level. As regards the former, we consider the concentration of the local credit market and the degree of localism of the banking system. Lacking data on the geographical distribution of bank deposits and loans, we measure concentration by the Herfindahl-Hirschman Index (*HHI*) computed on the number of branches in the province. The degree of localism is proxied by the share of branches held by Credit Cooperative Banks (*COOPERATIVE_BANKS*) in the province.[14] The expectation on the coefficient signs of these variables is uncertain. According to conventional wisdom, higher market concentration should go hand-in-hand with higher loan rates and a lower quantity of credit to borrowers. However, as the recent literature has pointed out,

[14] The inclusion of *COOPERATIVE_BANKS* aims to separate localism from functional distance. While a province with a banking system formed by only credit cooperatives banks has zero functional distance, it is also true that two banking systems with the same functional distance may show very different degrees of localism.

market concentration could also be beneficial for small firms due to the closer relationships with banks in a less competitive environment (Petersen and Rajan, 1995). In a similar vein, a high proportion of local banks in the credit market could increase credit availability to small local firms by reducing informational costs, but it might also be indicative of a closed and non-competitive banking system, less inclined to lend to young and riskier firms (Alessandrini and Zazzaro, 1999).

As regards institutional aspects, we consider the efficiency of courts in recovering bad loans measured by the logarithm of the average length (in days) of bankruptcy trials (*FAIL*).[15] Moreover, to take account of the unobserved specificity of Italian regions we add geographic dummies for the five Italian macro regions.[16] Finally, we add three time (wave) dummies.

4.1.2 Sample and Methodology

We look at the determinant of *Pr*(*RAT*) pooling the three MCC surveys. From the original sample, we excluded a number of observations due to missing data in balance-sheet entries, inconsistencies or extreme values. In particular, we excluded all firms which failed to indicate their headquarter location and which had an *R&D* value exceeding 100; firms with an *ROI* below the first and above the last percentile of its distribution; firms with a negative *DEBT* or with values greater than the last percentile. After this trimming process, our dataset decreased to 7,844 observations.

To address possible endogeneity and omitted variable problems in estimating the effect of banking variables (*OPERATIONAL_PROXIMITY* and *FUNCTIONAL_DISTACE* indexes, *HHI*, and *COOPERATIVE_BANKS*) on the likelihood of credit rationing, we also use instrumental variable techniques. In choosing instruments that may be correlated with banking development variables, but uncorrelated with the error term, we follow Guiso et al. (2004a) and exploit the fact that: (*i*) the number of branches in Italian regions and their distribution by size in 1936 were strictly regulated by the Bank of Italy and unrelated to regional economic development at that time; (*ii*) the geographical distribution of branches in 1936 remained substantially unaltered until the end of the 1970s and is significantly correlated with the local banking development in the 1990s. In this spirit, we construct a set of instrumental variables at the provincial level:[17] *OPERATIONAL_PROXIMITY_POPULATION* and

[15] The variable is available from ISTAT not at the provincial level, but only for larger Judiciary Districts (Corte d'Appello), each covering more than one province.

[16] Macro areas are: North-West, North-East, Centre, South and Island.

[17] The Guiso et al. (2004a) approach has been extended at the provincial level also by Benfratello et al. (2008) and Herrera and Minetti (2007). The provincial distribution of branches at the bank level is available from 1971 onwards.

OPERATIONAL_PROXIMITY_AREA in 1936 and 1971; *FUNCTIONAL_DISTANCE_KM* in 1971; *COOPERATIVE_BANKS* in 1936 and 1971; the share of branches owned by large banks, i.e., the former banks of national interest ("Istituti di Interesse Nazionale") and state-owned banks ("Istituti di Diritto Pubblico") in 1936 and 1971; and the value of *HHI* in 1971.

To take into account the possible endogeneity of *BANK_NUMBER* (Degryse and Ongena 2001), we add three more instruments proposed by Montoriol Garriga (2006) and Herrera and Minetti (2007): the ratio of bank debt to total assets at the industry level (*BANK_DEBT*); the average annual number of branches opened by entrants in a province over the period 1980–1985 (*ENTRANT_ΔBRANCHES_80-85*); the average annual number of branches opened net of branches closed by incumbents banks in a province over the period 1980–1985 (*INCUMBENT_ΔBRANCHES_80-85*). The first of such instrumental variables, (*BANK_DEBT*), captures the average degree of a firm's dependence on external finance due to exogenous industry-specific technological features. Therefore, it should be uncorrelated with the error term in Equation (1) and, to the extent that higher credit requirements necessitate more lenders, positively correlated with *BANK_NUMBER*. The rationale for the other two instruments lies in the effects that the numbers of competitors have on the opportunity to establish stable and exclusive lending relationships. In this vein, the opening of branches by incumbents in the past is expected to be negatively correlated with the number of bank relationships while, conversely, the *ENTRANT_ΔBRANCHES_80-85* should be positively correlated with *BANK_NUMBER*.

4.2 RESULTS

According to Table III, 15 per cent of firms in our sample claim to be rationed. This proportion decreases with firm size. Furthermore, while the average number of employees in rationed firms is 67, non-rationed firms have 110 employees. This difference is statistically significant at the 1 per cent level. Apart from size, rationed and non-rationed firms differ in a number of other features: firms facing credit constraints are, on average, significantly more indebted, less profitable and more likely linked to a bank located in the same local area. Furthermore, rationed firms are generally located in provinces where there are fewer cooperative banks (the one-tailed test for the *COOPERATIVE_BANKS* coefficient is significant at 6% level) and concentration is higher.

Finally, there are also significant differences with respect to our key distance variables. Rationed firms are usually located in provinces with a lower branch density (in terms of population) and greater functional distance.

Table V shows the results of probit and IV probit estimates of model (1) for the whole sample, while Table VI reports the results of IV probit for the two

Table III. Credit rationing: Univariate analysis on averages

The table reports the summary statistics of the explanatory variables of Equation (4) for rationed and non-rationed firms. The last column reports the p-values of a two-tailed t-test of the null hypothesis that the averages are equal in the two sub-samples of rationed and non-rationed firms. * significant at 10%; ** significant at 5%; *** significant at 1%. *RAT* is equal to one if the firm is credit-rationed and zero otherwise. Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j*, where the distance is proxied by the kilometric distance (*FUNCTIONAL_DISTANCE_KM*), the difference in social capital (*FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL*) or by a dissimilarity index in the economic structure (*FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE*), and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches ($\times$ 10,000) over surface area (*OPERATIONAL_PROXIMITY_AREA*). *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province. *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province. *BANK_NUMBER* is the number of banks with which the firm does business. *BANK_PROVINCE* is a dummy equal to one if the firm and its main bank are headquartered in the same province, and zero otherwise; *SIZE* is the logarithm of the number of workers; *R&D* is the ratio of those employed in R&D to total workers; *ROI* is the Return on Investment; *DEBT* is the logarithm of (1 + Debt-equity ratio); and *FAIL* is the logarithm of the average length, in days, of a bankruptcy trial, by judicial district.

	RAT = 1			*RAT* = 0			*t*-test (*p*-value)
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	
FUNCTIONAL_DISTANCE_KM	1,179	2.988	0.975	6,665	2.782	0.874	0.000***
FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL	1,179	0.989	0.495	6,665	0.865	0.413	0.000***
FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE	1,179	0.241	0.084	6,665	0.228	0.080	0.000***
FUNCTIONAL_DISTANCE_LATENT	1,179	2.281	0.872	6,665	2.078	0.761	0.000***
OPERATIONAL_PROXIMITY_POPULATION	1,179	5.350	1.655	6,665	5.807	1.440	0.000***
OPERATIONAL_PROXIMITY_AREA	1,179	23.675	28.600	6,665	24.622	29.07	0.3011
COOPERATIVE_BANKS	1,179	9.540	8.849	6,665	9.961	8.563	0.121
HHI	1,179	1.125	0.553	6,665	1.087	0.443	0.008***
BANK_NUMBER	1,179	6.120	3.867	6,665	6.134	4.015	0.914
BANK_PROVINCE	1,179	0.644	0.479	6,665	0.599	0.490	0.004***
SIZE	1,179	3.681	0.901	6,665	3.912	1.060	0.000***
R&D	1,179	0.766	2.696	6,665	0.675	2.612	0.275
ROI	1,179	10.343	6.75	6,665	12.962	7.605	0.000***
DEBT	1,179	6.223	0.995	6,665	5.820	0.986	0.000***
FAIL	1,179	7.821	0.184	6,665	7.830	0.161	0.074*

Table IV. Credit rationing (whole sample): First-stage regression results of IV Probit (Table V, column 5)

The table reports the OLS coefficients and, in brackets, the associated standard errors of the first stage regressions associated with the IV Probit estimates reported in column 5 of Table IV. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variables are the endogenous variables of Equation (1): *FUNCTIONAL_DISTANCE_KM* is the functional distance index, computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j*; *OPERATIONAL_PROXIMITY_POPULATION* is the operational proximity index, computed as the number of bank branches in the province per 10,000 inhabitants; *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province; *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province; *BANK_NUMBER* is the number of banks with which the firm does business. The set of instrumental variables includes a measure of *FUNCTIONAL_DISTANCE_KM* in 1971, *OPERATIONAL_PROXIMITY_POPULATION* in 1936 and 1971, *HHI* in 1971, *COOPERATIVE_BANKS* in 1936 and 1971, the share of branches held by large banks (*SHARE_LARGE_BANKS*) in 1936 and 1971, the average number of new branches held by incumbent and entrant banks in the province over the period 1980–1985 (*INCUMBENT_ΔBRANCHES_80–85* and *ENTRANT_ΔBRANCHES_80–85*), the average ratio of bank debt over total asset at industry level (Ateco, 2 digit) over the period 1995–2003 (*BANK_DEBT*). The other exogenous controls are: *BANK_PROVINCE* is a dummy equal to one if the firm and its main bank are headquartered in the same province, and zero otherwise; *SIZE* is the logarithm of the number of workers; *R&D* is the ratio of those employed in R&D to total workers; *ROI* is the Return on Investment; *DEBT* is the logarithm of (1 + Debt-equity ratio); and *FAIL* is the logarithm of the average length, in days, of a bankruptcy trial, by judicial district. All regressions include the constant and wave, geographic (five macro-areas) and industry (Pavitt classification) dummies. As diagnostic, the table reports the Adjusted-R^2, the p-values for the F-tests that the Instrumental Variables, the Pavitt dummies and the Geographic dummies are jointly equal to zero.

	Endogenous variables				
	FUNCTIONAL_DISTANCE KM	*OPERATIONAL_PROXIMITY_ POPULATION*	*COOPERATIVE_ BANKS*	*HHI*	*BANK_NUMBER*
Instrumental variables					
COOPERATIVE_BANKS_1936	0.880***	−0.966***	−19.036***	−0.365***	0.118
	[0.098]	[0.093]	[0.944]	[0.029]	[0.615]
COOPERATIVE_BANKS_1971	−0.368***	0.264***	3.092***	0.105***	−0.142
	[0.037]	[0.035]	[0.351]	[0.011]	[0.229]
SHARE_LARGE_BANKS_1936	0.259***	−1.284***	−9.930***	−0.338***	−1.285**
	[0.100]	[0.094]	[0.955]	[0.030]	[0.622]
SHARE_LARGE_BANKS_1971	−1.325***	0.425***	−10.074***	0.919***	1.415**
	[0.115]	[0.109]	[1.106]	[0.034]	[0.721]
OPERATIONAL_PROXIMITY_POPULATION_1936	−0.051***	−0.040***	3.118***	−0.054***	−0.176**
	[0.012]	[0.012]	[0.120]	[0.004]	[0.078]

Table IV. (Continued)

	Endogenous variables				
	FUNCTIONAL_DISTANCE KM	*OPERATIONAL_PROXIMITY_ POPULATION*	*COOPERATIVE_ BANKS*	*HHI*	*BANK_NUMBER*
OPERATIONAL_PROXIMITY_POPULATION_1971	−0.453***	1.008***	2.526***	0.074***	0.362***
	[0.017]	[0.016]	[0.162]	[0.005]	[0.106]
HHI_1971	−0.224***	0.180***	−5.556***	0.642***	−0.387***
	[0.023]	[0.022]	[0.218]	[0.007]	[0.142]
FUNCTIONAL_DISTANCE_KM_1971	0.397***	−0.123***	−1.550***	0.035***	−0.098
	[0.010]	[0.010]	[0.099]	[0.003]	[0.064]
BANK_DEBT	0.365	0.566***	0.373	−0.100	9.161***
	[0.224]	[0.213]	[2.151]	[0.067]	[1.401]
INCUMBENT_ΔBRANCHES_80–85	−0.086***	0.073***	1.526***	−0.012***	−0.070**
	[0.005]	[0.005]	[0.052]	[0.002]	[0.034]
ENTRANT_ΔBRANCHES_80–85	−0.078***	0.039***	−1.545***	−0.001	−0.030
	[0.004]	[0.004]	[0.041]	[0.001]	[0.027]
Exogenous variables					
BANK_PROVINCE	−0.079***	−0.006	0.172	0.006	−0.301***
	[0.013]	[0.012]	[0.122]	[0.004]	[0.080]
SIZE	−0.005	0.007	0.163***	0.004**	1.826***
	[0.006]	[0.006]	[0.059]	[0.002]	[0.039]
R&D	0.000	−0.001	0.012	−0.000	0.030**
	[0.002]	[0.002]	[0.023]	[0.001]	[0.015]
ROI	−0.002**	0.001	0.015*	0.000	−0.037***
	[0.001]	[0.001]	[0.008]	[0.000]	[0.005]
DEBT	0.003	0.006	0.150**	−0.002	0.534***
	[0.007]	[0.006]	[0.063]	[0.002]	[0.041]
FAIL	0.385***	0.480***	−4.653***	−0.479***	1.711***
	[0.057]	[0.054]	[0.547]	[0.017]	[0.356]
Observations	7,844	7,844	7,844	7,844	7,844
Adjusted-R^2	0.633	0.880	0.636	0.806	0.283
F-test on IV	0.000	0.000	0.000	0.000	0.000
F-test on Pavitt dummies	0.014	0.219	0.000	0.498	0.000
F-test on Geographic dummies	0.000	0.000	0.000	0.000	0.000

Table V. Credit rationing (whole sample): Probit estimation of Equation (1)

The table reports regression coefficients and, in brackets, the associated standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable (*RAT*) is equal to one if the firm is credit-rationed and zero otherwise. Columns 1–4 report the estimates of the standard probit model, while columns 5–8 refer to the Instrumental Variable (IV) estimates. To control for endogeneity of the distance and banking variables, all IV regressions are estimated with Newey's two-step estimator (IVPROBIT using Stata 10 SE package). Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j* (*FUNCTIONAL_DISTANCE_KM*) and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches ($\times$ 10,000) over surface area (*OPERATIONAL_PROXIMITY_AREA*). *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province. *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province. *BANK_NUMBER* is the number of banks with which the firm does business. In columns 5–8, all these variables are taken as endogenous and, as instruments for these variables, we include a measure of *FUNCTIONAL_DISTANCE_KM* in 1971, *OPERATIONAL_PROXIMITY_POPULATION* (columns 5–6) and *OPERATIONAL_PROXIMITY_AREA* (columns 7–8) in 1936 and 1971, *HHI* in 1971, *COOPERATIVE_BANKS* in 1936 and 1971, the share of branches held by large banks in 1936 and 1971, the average number of new branches held by incumbent and entrant banks in the province over the period 1980–1985, the average ratio of bank debt over total asset at industry level (Ateco, 2 digit) over the period 1995–2003. The other exogenous controls are: *BANK_PROVINCE* is a dummy equal to one if the firm and its main bank are headquartered in the same province, and zero otherwise; *SIZE* is the logarithm of the number of workers; *R&D* is the ratio of those employed in R&D to total workers; *ROI* is the Return on Investment; *DEBT* is the logarithm of (1 + Debt-equity ratio); and *FAIL* is the logarithm of the average length, in days, of a bankruptcy trial, by judicial district. All regressions include the constant and wave, geographic (five macro-areas) and industry (Pavitt classification) dummies. As diagnostic, the table reports the p-values of the Wald χ^2-statistic for the likelihood ratio test of the goodness of fit of the regression, of the test for over-identifying restrictions (OIR, the null is the validity of the instrument set), and of the F-test of the null hypothesis that the geographic dummies are jointly equal to zero. The bottom lines report the estimated change in the probability of *RAT* due to a change in the *FUNCTIONAL_DISTANCE* indexes: (1) from the first to the third quartile of the sample distribution, and (2) from their actual provincial values in 1995 to those in 2003, both for the representative firm located in the Centre-North and in the South.

Dep Var: Pr(*RAT*)	Probit				IV Probit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Banking distance variables								
FUNCTIONAL_DISTANCE_KM	0.045*		0.068**		0.169***		0.201***	
	[0.024]		[0.027]		[0.057]		[0.061]	
FUNCTIONAL_DISTANCE_LATENT		0.063**		0.090***		0.197***		0.230***
		[0.029]		[0.032]		[0.064]		[0.068]
OPERATIONAL_PROXIMITY_POPULATION	−0.064**	−0.062**			0.046	0.026		
	[0.026]	[0.026]			[0.072]	[0.067]		
OPERATIONAL_PROXIMITY_AREA			0.001	0.001			0.003*	0.002
			[0.001]	[0.001]			[0.002]	[0.002]

Table V. (Continued)

Dep Var: Pr(*RAT*)	Probit				IV Probit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Other banking variables								
COOPERATIVE_BANKS	0.005*	0.005*	0.003	0.003	0.002	0.004	0.011*	0.010*
	[0.003]	[0.003]	[0.003]	[0.003]	[0.006]	[0.006]	[0.006]	[0.006]
HHI	−0.018	−0.019	0.004	0.001	−0.121	−0.116	0.09	0.049
	[0.053]	[0.053]	[0.072]	[0.071]	[0.076]	[0.075]	[0.122]	[0.114]
BANK_NUMBER	0.015***	0.015***	0.014***	0.014***	−0.112*	−0.106*	−0.067	−0.07
	[0.005]	[0.005]	[0.005]	[0.005]	[0.063]	[0.062]	[0.054]	[0.054]
BANK_PROVINCE	0.115***	0.117***	0.115***	0.117***	0.093**	0.095**	0.108**	0.107**
	[0.037]	[0.037]	[0.037]	[0.037]	[0.044]	[0.044]	[0.043]	[0.043]
Other control variables								
SIZE	−0.111***	−0.110***	−0.109***	−0.108***	0.122	0.111	0.04	0.046
	[0.022]	[0.022]	[0.022]	[0.022]	[0.116]	[0.114]	[0.099]	[0.099]
R&D	0.014**	0.014**	0.014**	0.014**	0.018***	0.017***	0.017**	0.017**
	[0.006]	[0.006]	[0.006]	[0.006]	[0.007]	[0.007]	[0.006]	[0.006]
ROI	−0.016***	−0.016***	−0.016***	−0.016***	−0.021***	−0.021***	−0.019***	−0.019***
	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]	[0.004]	[0.004]
DEBT	0.213***	0.213***	0.213***	0.214***	0.281***	0.279***	0.257***	0.260***
	[0.019]	[0.019]	[0.019]	[0.019]	[0.039]	[0.039]	[0.035]	[0.035]
FAIL	−0.379**	−0.368**	−0.419**	−0.401**	−0.351*	−0.286	−0.221	−0.189
	[0.183]	[0.181]	[0.189]	[0.188]	[0.186]	[0.181]	[0.189]	[0.192]
Observations	7844	7844	7844	7844	7844	7844	7844	7844
Pseudo-R^2	0.067	0.067	0.066	0.067				
Wald test					0.000	0.000	0.000	0.000
OIR test					0.387	0.452	0.398	0.498
F-test on Geographic dummies	0.004	0.008	0.000	0.000	0.106	0.147	0.002	0.016
Estimated effect on Pr(*RAT*) *due to a change in FUNCTIONAL_DISTANCE*:								
from 25° to 75° percentile, whole sample	1.290	1.410	1.960	2.010	4.870	4.400	5.750	5.100
from 1996 to 2003, centre-northern sample	0.351	0.389	0.535	0.555	1.340	1.240	1.580	1.430
from 1996 to 2003, southern sample	0.919	1.230	1.420	1.780	3.540	3.930	4.370	4.740

Table VI. Credit rationing (small vs large firms): IV Probit estimation of Equation (1)

The table reports regression coefficients and, in brackets, the associated standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%. *Small* is the sub-sample of firms employing 11–50 workers, while *Large* refer to firms with more than 50 employees. The dependent variable (*RAT*) is equal to one if the firm is credit-rationed and zero otherwise. To control for endogeneity of the distance and banking variables, all probit regressions are estimated with Newey's two-step estimator (IVPROBIT using Stata 10 SE package). Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j* (*FUNCTIONAL_DISTANCE_KM*) and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches (× 10,000) over surface (*OPERATIONAL_PROXIMITY_AREA*). *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province. *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province. *BANK_NUMBER* is the number of banks with which the firm does business. All these variables are taken as endogenous. As instruments for these variables, we include a measure of *FUNCTIONAL_DISTANCE_KM* in 1971, *OPERATIONAL_PROXIMITY_POPULATION* (columns 1–4) and *OPERATIONAL_PROXIMITY_AREA* (columns 5–8) in 1936 and 1971, *HHI* in 1971, *COOPERATIVE_BANKS* in 1936 and 1971, the share of branches held by large banks in 1936 and 1971, the average number of new branches held by incumbent and entrant banks in the province over the period 1980–1985, the average ratio of bank debt over total assets at industry level (Ateco, 2 digit) over the period 1995–2003. The other exogenous controls are: *BANK_PROVINCE* is a dummy equal to one if the firm and its main bank are headquartered in the same province, and zero otherwise; *SIZE* is the logarithm of the number of workers; *R&D* is the ratio of employed in R&D activities on total workers; *ROI* is the Return on Investment; *DEBT* is the logarithm of (1 + Debt-equity ratio); and *FAIL* is the logarithm of the average length, in days, of a bankruptcy trial, by judicial district. All regressions include the constant and wave, geographic (five macro-areas) and industry (Pavitt classification) dummies. As diagnostic, the table reports the p-values of the Wald χ^2-statistic for the likelihood ratio test of the goodness of fit of the regression, of the test for over-identifying restrictions (OIR, the null is the validity of the instrument set), and of the F-test of the null hypothesis that the geographic dummies are jointly equal to zero. The bottom lines report the estimated change in the probability of *RAT* due to a change in the *FUNCTIONAL_DISTANCE* indexes: (1) from the first to the third quartile of the sample distribution, and (2) from their actual provincial values in 1995 to those in 2003, both for the representative firm located in the Centre-North and in the South. § indicates that the null hypothesis that the estimated change in the probability of *RAT* is equal for small and large firms is rejected at a 2.5% level of significance under a two-tailed *t*-test.

Dep Var: Pr(*RAT*)	(1) Small	(2) Large	(3) Small	(4) Large	(5) Small	(6) Large	(7) Small	(8) Large
Banking distance variables								
FUNCTIONAL_DISTANCE_KM	0.211**	0.166*			0.212**	0.217**		
	[0.083]	[0.085]			[0.084]	[0.098]		
FUNCTIONAL_DISTANCE_LATENT			0.243***	0.190**			0.240***	0.240**
			[0.092]	[0.096]			[0.093]	[0.107]

Table VI. (Continued)

Dep Var: Pr(*RAT*)	(1) Small	(2) Large	(3) Small	(4) Large	(5) Small	(6) Large	(7) Small	(8) Large
OPERATIONAL_PROXIMITY_POPULATION	0.134	−0.045	0.107	−0.067				
	[0.106]	[0.101]	[0.096]	[0.095]				
OPERATIONAL_PROXIMITY_AREA					0.002	0.003	0.002	0.002
					[0.002]	[0.003]	[0.002]	[0.003]
Other banking variables								
COOPERATIVE_BANKS	−0.004	0.008	−0.003	0.01	0.011	0.009	0.01	0.007
	[0.008]	[0.009]	[0.008]	[0.008]	[0.008]	[0.010]	[0.008]	[0.009]
HHI	−0.168	−0.061	−0.162	−0.054	0.051	0.096	0.009	0.046
	[0.103]	[0.122]	[0.101]	[0.121]	[0.161]	[0.208]	[0.150]	[0.195]
BANK_NUMBER	−0.232**	0.06	−0.217**	0.059	−0.191**	0.082	−0.187**	0.073
	[0.098]	[0.069]	[0.094]	[0.069]	[0.085]	[0.065]	[0.085]	[0.064]
BANK_PROVINCE	0.071	0.157**	0.075	0.155**	0.082	0.168**	0.083	0.164**
	[0.059]	[0.067]	[0.058]	[0.067]	[0.058]	[0.068]	[0.057]	[0.067]
Other control variables								
SIZE	0.274*	−0.271**	0.253*	−0.270**	0.206	−0.313**	0.202	−0.299**
	[0.152]	[0.135]	[0.146]	[0.135]	[0.134]	[0.129]	[0.133]	[0.128]
R&D	0.017*	0.011	0.017*	0.011	0.017**	0.01	0.017**	0.011
	[0.009]	[0.011]	[0.009]	[0.011]	[0.009]	[0.010]	[0.009]	[0.010]
ROI	−0.021***	−0.016***	−0.021***	−0.016***	−0.020***	−0.015**	−0.020***	−0.015**
	[0.004]	[0.006]	[0.004]	[0.006]	[0.004]	[0.006]	[0.004]	[0.006]
DEBT	0.317***	0.184***	0.311***	0.187***	0.299***	0.164**	0.298***	0.173***
	[0.047]	[0.067]	[0.046]	[0.067]	[0.042]	[0.065]	[0.042]	[0.065]
FAIL	−0.552**	−0.344	−0.466**	−0.275	−0.28	−0.342	−0.248	−0.313
	[0.246]	[0.298]	[0.236]	[0.292]	[0.242]	[0.323]	[0.245]	[0.326]
Observations	4674	3170	4674	3170	4674	3170	4674	3170
Wald test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
OIR test	0.394	0.957	0.430	0.960	0.345	0.790	0.379	0.788
F-test on Geographic dummies	0.044	0.759	0.049	0.837	0.001	0.089	0.004	0.293
Estimated effect on Pr(*RAT*) *due to a change in FUNCTIONAL_DISTANCE*:								
from 25° to 75° percentile, whole sample	6.680§	4.130	6.420§	3.580	6.700§	5.420	6.360§	4.520
from 1996 to 2003, centre-northern sample	1.810§	1.140	1.640§	1.040	1.810§	1.510	1.620§	1.320
from 1996 to 2003, southern sample	4.330§	3.360	4.760§	3.660	4.480	4.470	4.810	4.720

subsamples of small and medium-large firms (i.e., firms with less and more than 50 employees, respectively). In all specifications, the Wald test is significant and the over-identification test supports the validity of the instruments. In addition, first-stage regression results, reported in Table IV, show that, once net of the effect of the other explanatory variables, instruments are generally significantly correlated with the endogenous variables and the coefficients have the expected signs. In particular, the banking market structure prior to deregulation is a significant predictor of current structure, while firms' dependence on external finance and the opening of banks by incumbents significantly affect multiple lending, respectively increasing and reducing the number of banks the firm operates with. Finally, the F-test always rejects the hypothesis that instruments are jointly insignificant and the Adjusted-R^2 is satisfactorily high.[18]

To assess the differential impact of functional distance on credit availability, we report in the last three rows of Tables V and VI the change in the probability of credit rationing for the average firm and for the representative firms headquartered in centre-northern and southern provinces associated with, respectively, the theoretical increase in *FUNCTIONAL_DISTANCE* indexes from the first to the third quartile of its sample distribution, and with the observed increase from 1996 to 2003 in the Centre-North and in the South. Our expectation is that credit availability to small firms, whose borrowing relies greatly upon soft, proprietary information, is more negatively affected by functional distance. In the same vein, we expect that in southern regions, where bank consolidation process has led to the almost complete disappearance of an autonomous banking system and where the economic environment is extremely weak, the increase in functional distance has exacerbated financing constraints to local firms.

4.2.1 Banking Distance Variables

Overall, what emerges from the regression analysis is that functional distance makes access to credit significantly harder for local firms, while operational proximity, though somewhat negatively correlated with credit rationing, is not significantly different from zero once the endogeneity issue is taken into account.

The sign of the *FUNCTIONAL_DISTANCE* coefficients is positive both in the probit and IV probit estimates and its effect on credit rationing is larger when operational proximity is measured with respect to province size. The statistical and economic significance of *FUNCTIONAL_DISTANCE_KM* and *FUNCTIONAL_DISTANCE_LATENT* greatly increase after instrumenting these variables. For example, if we consider the probit estimates the change experienced

[18] For the sake of space, in Table IV we report only results of first-stage regressions relative to the whole sample basic specification (Table V, column 5). All the other first-stage regressions provide virtually identical results and are available on request.

by functional distance in Italy from 1996 to 2003 led to an increase in the likelihood of credit rationing for the centre-northern (southern) firm that varies from 0.35 to 0.56 per cent (from 0.92 to 1.78 per cent) according to the operational proximity index used. These numbers rise by more than a factor of three in IV estimates, arriving at a sizeable 4.74 per cent for the average southern firm (Table V, column 8). Operational proximity has the expected beneficial effect on the probability of credit rationing only in probit estimates when the number of branches is scaled by population. In all other cases, the *OPERATIONAL_PROXIMITY* coefficient is statistically insignificant.

When we split our sample into small and large firms, the adverse effect of functional distance is, statistically and economically, stronger for small firms. To illustrate, a small firm in a province with a local banking system having *FUNCTIONAL_DISTANCE_KM* within the first quartile of its distribution has about 7 per cent less likelihood of being rationed than a small firm in a province with *FUNCTIONAL_DISTANCE_KM* in the third quartile; for large firms this effect is slightly above 4 per cent (Table VI, columns 1 and 2). If we focus on the actual increase in functional distance from 1996 to 2003, the impact of this change on *Pr*(*RAT*) was higher for southern than centre-northern firms: 4.3 per cent versus 1.8 per cent for the representative small firm and 3.4 per cent versus 1.1 per cent for the representative large firm, respectively[19]. This differential impact of functional distance across small and large firms is statistically significant at the 1 per cent level, except for southern firms when operational proximity is scaled by province area. A possible explanation for this relatively small difference between small and large borrowers in the South lies in the differences between the average large (small) centre-northern and southern firms, which drive the magnitude of the marginal impact of *FUNCTIONAL_DISTANCE* on *Pr(RAT)*. For example, the average size of large southern firms is smaller than that of large average firms (174 versus 224 employees, respectively), whereas the average size of southern small firms is slightly higher (27 versus 26 employees). Thus, these differences shrink the gap between small and large firms in the sample of southern firms, relative to the sample of centre-northern firms. By same token, *R&D*, whose estimated coefficient is positive, is more than twice as high in large centre-northern than in large southern firms, whereas this difference decreases to around 70 per cent for small firms located in the two areas. By contrast, the geographical difference in variables reducing the probability of credit rationing, like *ROI*, are relatively higher for small than for large firms.

The positive correlation between *Pr*(*RAT*) and *FUNCTIONAL_DISTANCE* is confirmed when we weight the latter by cultural and economic distance indicators.

[19] For small (large) firms, these numbers are computed as the mean of the estimated effects on *Pr(RAT)* due to a change in *FUNCTIONAL_DISTANCE* reported in the odd (even) columns of Table VI.

In these specifications, the differential effect between small and large firms is still greater and always statistically significant at the 2.5 per cent level (Table VII).

On the whole, regression results are broadly consistent with the theoretical prediction by which larger firms suffer less from the lack of banks' decisional centres in the province where they are located.[20] However, they also suggest that *FUNCTIONAL_DISTANCE* is especially harmful in peripheral regions, where large firms often show economic weakness, financial fragility and information opaqueness similar to those of small firms.

4.2.2 Banking Control Variables

Moving on to the results with banking control variables, the probit estimates show that the larger the number of banks with which a firm does business the higher is the probability of credit rationing (Table V, columns 1–4). This evidence is consistent with previous findings for Italy by Angelini et al. (1998) and Guelpa and Tirri (2006), who, however, do not allow for the endogeneity of multiple lending relationships. By contrast, after instrumenting, *BANK_NUMBER* has a negative impact on *Pr*(*RAT*), even if it is not always significant. When we split our sample according to firm size (Table VI), we find that multiple lending has the effect of increasing credit availability to small firms, while for large firms the effect is statistically not different from zero. Such diversity suggests that informational capture by the main bank is a concrete risk for small firms, from which they can escape by borrowing from other banks. We also find that having the main bank headquartered in the same province (*BANK_PROVINCE*) raises difficulties in securing adequate amounts of finance. This result holds no matter what estimation technique is used, but it appears to be driven especially by large firms. It suggests that relationship loans may be harmful to firms when the lender is a local bank that may informationally capture their customers.

The adverse impact of a banking system with a large share of local banks on the probability of being credit-rationed is partly confirmed by the sign of the *COOPERATIVE_BANKS* coefficient. For the whole sample, it is positive and significant in the probit estimates with *OPERATIONAL_PROXIMITY_POPULATION* and in the IV probit estimates with *OPERATIONAL_PROXIMITY_AREA*, and insignificant in other cases.

With regard to market concentration, the estimated coefficient for *HHI* is never statistically different from zero at conventional confidence levels, even if for small firms its impact is beneficial at a significance slightly above the 10 per cent level (Table VI, columns 1 and 3).

[20] In the working paper version of this paper we also estimate a probit model for the whole sample including an interaction term between functional distance and firm size and find that the sign of this regressor is significantly negative (Alessandrini et al., 2006).

Table VII. Credit rationing (small vs large firms): IV Probit estimation of Equation (1)

The table reports regression coefficients and, in brackets, the associated standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%. *Small* is the sub-sample of firms employing 11–50 workers, while *Large* refer to firms with more than 50 employees. The dependent variable (*RAT*) is equal to one if the firm is credit-rationed and zero otherwise. To control for endogeneity of the distance and banking variables, all probit regressions are estimated with Newey's two-step estimator (IVPROBIT using Stata 10 SE package). Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j*, where the distance is proxied by the difference in social capital (*FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL*) or by a dissimilarity index in the economic structure (*FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE*). The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches ($\times$ 10,000) over surface (*OPERATIONAL_PROXIMITY_AREA*). *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province. *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province. *BANK_NUMBER* is the number of banks with which the firm does business. All these variables are taken as endogenous. As instruments for these variables, we include a measure of *FUNCTIONAL_DISTANCE_KM* in 1971, *OPERATIONAL_PROXIMITY_POPULATION* (columns 1–4) and *OPERATIONAL_PROXIMITY_AREA* (columns 5–8) in 1936 and 1971, *HHI* in 1971, *COOPERATIVE_BANKS* in 1936 and 1971, the share of branches held by large banks in 1936 and 1971, the average number of new branches held by incumbent and entrant banks in the province over the period 1980–1985, the average ratio of bank debt over total assets at industry level (Ateco, 2 digit) over the period 1995–2003. The other exogenous controls are: *BANK_PROVINCE* is a dummy equal to one if the firm and its main bank are headquartered in the same province, and zero otherwise; *SIZE* is the logarithm of the number of workers; *R&D* is the ratio of employed in R&D activities on total workers; *ROI* is the Return on Investment; *DEBT* is the logarithm of (1 + Debt-equity ratio); and *FAIL* is the logarithm of the average length, in days, of a bankruptcy trial, by judicial district. All regressions include the constant and wave, geographic (five macro-areas) and industry (Pavitt classification) dummies. As diagnostic, the table reports the p-values of the Wald χ^2-statistic for the likelihood ratio test of the goodness of fit of the regression, of the test for over-identifying restrictions (OIR, the null is the validity of the instrument set), and of the F-test of the null hypothesis that the geographic dummies are jointly equal to zero. The bottom lines report the estimated change in the probability of *RAT* due to a change in the *FUNCTIONAL_DISTANCE* indexes: (1) from the first to the third quartile of the sample distribution, and (2) from their actual provincial values in 1995 to those in 2003, both for the representative firm located in the Centre-North and in the South. indicates that the null hypothesis that the estimated change in the probability of *RAT* is equal for small and large firms is rejected at a 2.5% level of significance under a two-tailed t-test.

Dep Var: Pr(*RAT*)	(1) Small	(2) Large	(3) Small	(4) Large	(5) Small	(6) Large	(7) Small	(8) Large
Banking distance variables								
FUNCTIONAL_DISTANCE_SOCIAL_CAPITAL	0.482***	0.379*			0.465***	0.449**		
	[0.175]	[0.198]			[0.177]	[0.205]		
FUNCTIONAL_DISTANCE_ECONOMIC_STRUCTURE			1.968**	1.654*			2.184**	1.844*
			[0.863]	[0.930]			[0.946]	[1.065]

OPERATIONAL_PROXIMITY_POPULATION	0.028	−0.131	0.057	−0.088				
	[0.078]	[0.083]	[0.087]	[0.093]				
OPERATIONAL_PROXIMITY_AREA					0.000	0.000	0.001	0.001
					[0.002]	[0.002]	[0.002]	[0.002]
Other banking variables								
COOPERATIVE_BANKS	0.002	0.013	0.001	0.012	0.005	0.002	0.009	0.004
	[0.007]	[0.008]	[0.007]	[0.008]	[0.007]	[0.008]	[0.008]	[0.009]
HHI	−0.131	−0.029	−0.154	−0.055	−0.101	−0.075	−0.053	−0.043
	[0.095]	[0.119]	[0.100]	[0.123]	[0.132]	[0.177]	[0.140]	[0.188]
BANK_NUMBER	−0.178**	0.058	−0.165*	0.068	−0.183**	0.053	−0.165*	0.082
	[0.087]	[0.069]	[0.087]	[0.069]	[0.085]	[0.065]	[0.084]	[0.065]
BANK_PROVINCE	0.082	0.145**	0.087	0.161**	0.080	0.147**	0.091	0.168**
	[0.057]	[0.066]	[0.057]	[0.068]	[0.057]	[0.066]	[0.058]	[0.069]
Other control variables								
SIZE	0.196	−0.269**	0.182	−0.281**	0.200	−0.266**	0.177	−0.310**
	[0.137]	[0.135]	[0.136]	[0.135]	[0.133]	[0.129]	[0.133]	[0.129]
R&D	0.017*	0.010	0.017*	0.010	0.017*	0.011	0.017**	0.011
	[0.009]	[0.011]	[0.009]	[0.011]	[0.009]	[0.010]	[0.009]	[0.010]
ROI	−0.020***	−0.016***	−0.020***	−0.016***	−0.020***	−0.017***	−0.020***	−0.015**
	[0.004]	[0.006]	[0.004]	[0.006]	[0.004]	[0.006]	[0.004]	[0.006]
DEBT	0.295***	0.190***	0.288***	0.183***	0.297***	0.193***	0.287***	0.171***
	[0.043]	[0.067]	[0.042]	[0.067]	[0.042]	[0.065]	[0.042]	[0.065]
FAIL	−0.296	−0.138	−0.252	−0.120	−0.243	−0.300	−0.091	−0.238
	[0.229]	[0.292]	[0.232]	[0.296]	[0.243]	[0.328]	[0.282]	[0.359]
Observations	4,674	3,170	4,674	3,170	4,674	3,170	4,674	3,170
Wald test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
OIR test	0.438	0.935	0.206	0.892	0.398	0.739	0.239	0.530
F-test on Geographic dummies	0.049	0.644	0.034	0.374	0.002	0.327	0.000	0.003
Estimated effect on Pr(*RAT*) *due to a change in FUNCTIONAL_DISTANCE*:								
from 25° to 75° percentile, whole sample	5.060§	2.950	4.830§	3.300	4.890§	3.470	5.350§	3.680
from 1996 to 2003, centre-northern sample	0.985§	0.637	0.973§	0.666	0.952§	0.755	1.080§	0.748
from 1996 to 2003, southern sample	5.590§	4.250	2.410§	1.880	5.360§	5.190	2.680§	2.100

4.2.3 Other Control Variables

Looking at firm-specific characteristics, we find, as expected, that more profitable firms are less likely to be rationed, while those with a large R&D department and with a high degree of leverage encounter more difficulties in accessing bank credit. These effects are statistically highly significant and, *ceteris paribus*, stronger for small firms. The impact of *SIZE* on the probability of credit rationing is negative only in probit estimation. After instrumenting banking variables, the effect of *SIZE* is statistically nil on average, but is qualitatively different and significant in our two subsamples. For large firms, an increase in the number of employees is associated with a decrease in the likelihood of being rationed. Conversely, within the sample of small firms, it is the smallest ones which are less likely to be rationed. Arguably this finding might reflect the fact that the credit requested by micro firms is usually very small for banks that consequently prefer to reject unworthy borrowers instead of negotiating a reduction in the amount of loan granted.

Concerning geographical unobserved effects, we find that only South and Islands dummies are statistically significant. In particular, firms located in southern regions face a 6–10 per cent higher probability of being rationed than those in the North; this effect is even greater (between 10 and 16 per cent) for firms in Sicily and Sardinia. Finally, the efficiency of courts in recovering bad loans increases the rationing probability even if with IV probit estimations this effect is statistically significant only for small firms. Albeit unexpected, this finding is consistent with a sort of "laziness effect" (Manove et al., 2001; Zazzaro, 2005): where the enforcement of credit contracts by courts is stricter banks find it profitable to reduce their costly screening effort; this increases the probability of evaluating borrowers erroneously and rationing credit to worthy borrowers.

5. Investment-Cash Flow Sensitivity

5.1 THE ECONOMETRIC MODEL

Our second econometric exercise is aimed at testing the impact of distance on the sensitivity of investment to cash flow. In particular, we estimate a dynamic panel investment model in which we include cash flow interacting with operational proximity and functional distance.[21]

[21] We also consider a possible direct effect of distance on investment by introducing *FUNCTIONAL_DISTANCE* and *OPERATIONAL_PROXIMITY* indexes as isolated terms in equation (2). However, since they prove to be statistically insignificant without affecting the other estimation coefficients, we prefer to follow our topic strictly and limit analysis to the indirect effect of distance on financial constraints for firm investment.

Data on capital stock and cash flow come from firm balance-sheets. Merging the three waves, we obtain data over the period 1996–2003 for 526 firms. Furthermore, to get a balanced panel we drop 245 firms and end up with a sample of 279 firms for a total of 2,232 observations.

The basic investment equation we estimate is:

$$
\begin{aligned}
\left(\frac{INVESTMENT_t}{K_{t-1}}\right)_{ij} = \alpha + \beta_1 \left(\frac{INVESTMENT_{t-1}}{K_{t-2}}\right)_{ij} + \beta_2 \left(\frac{CASH\_FLOW_t}{K_{t-1}}\right)_{ij} \\
+ \beta_3 \left(\frac{CASH\_FLOW_t}{K_{t-1}}\right)_{ij} \times OPERATIONAL\_PROXIMITY_{jt} \\
+ \beta_4 \left(\frac{CASH\_FLOW_t}{K_{t-1}}\right)_{ij} \times FUNCTIONAL\_DISTANCE_{jt} \\
+ \beta_5 GROWTH_{ijt-1} + \eta_i + \nu_t + \varepsilon_{i,t} \qquad (2)
\end{aligned}
$$

where the subscripts refer to the *i*-th firm, located in province *j* at time *t*, *K* is the capital stock at the end of the fiscal year, computed as the tangible and intangible assets gross of depreciation allowances; *INVESTMENT* is defined as the variation between time *t* and *t-1* of the firm's capital stock; *CASH_FLOW* is defined as net income plus depreciation allowances; *GROWTH* is the growth rate of sales which is a proxy of firm profitability; η_i are time-invariant individual effects; ν_t is the time-specific effect, and ε_{it} is an idiosyncratic shock. Finally, in the basic specification of Equation (2) we include Pavitt's industry classification dummies to control for firm specificity, and geographic dummies to control for other unobserved local fixed effects.

We expect a financially constrained firm to exhibit a positive correlation between cash flow and investment and consider as a measure of financing constraints the marginal effect of cash flow on investment, $\partial(INVESTMENT/K)/\partial(CASH\_FLOW/K)$. Given the inclusion of the two interaction terms between distances and cash flow, the sensitivity of investment to cash flow becomes:

$$
\begin{aligned}
\frac{\partial(INVESTMENT/K)}{\partial(CASH\_FLOW/K)} = \beta_2 + \beta_3 OPERATIONAL\_PROXIMITY \\
+ \beta_4 FUNCTIONAL\_DISTANCE \qquad (3)
\end{aligned}
$$

Thus, according to $\beta_3 \gtrless 0$ and $\beta_4 \gtrless 0$, operational proximity and functional distances may increase or reduce firms' financing constraints. We also consider specifications with the additional interaction term *CASH_FLOW×FUNCTIONAL_DISTANCE×SIZE* to check whether the effect of functional distance on financing constraints is stronger for small than for large firms, where firm size is measured by classes of employees to allow for possible non-linearities.

Due to the dynamic structure of the model and the short time dimension of the panel, the LSDV estimator is biased and inconsistent (Nickell, 1981). For that reason, and to control for endogeneity and omitted variable bias, we use the Generalized Method of Moments (GMM) and, in particular, the System-GMM (Blundell and Bond, 1998).

5.2 RESULTS

To interpret a positive (negative) impact of distance indexes on (3) as strengthening (weakening) financing constraints for firms we have to assume that the sensitivity of investment to cash flow is a reliable measure of financing constraints. This assumption has been challenged by Kaplan and Zingales (1997), who present evidence that investment by financially constrained firms may exhibit significantly lower sensitivity to cash flow than investment by non-financially constrained firms. To address this important criticism, we estimate Equation (2) excluding the interaction terms and discriminating rationed from non-rationed firms on the ground of the variable *RAT*. The estimation results show that rationed firms report a greater elasticity of investment with respect to cash flow than non-rationed ones, suggesting that $\partial(INVESTMENT/K)/\partial(CASH\_FLOW/K)$ is a good proxy for financing constraints.[22]

On the whole, the results on the sensitivity of investment to cash flow are consistent with those on credit rationing described in the previous section. Investment of firms in provinces with a functionally more distant banking system is more sensitive to cash flow, and this effect decreases with firm size.

Table VIII displays the estimation results of our basic (columns 1 to 4) and augmented (columns 5 to 8) specifications. For each regression, the over-identification test supports the validity of the instrument set at 5% and the Arellano and Bond (1991) autocorrelation tests show the expected values at the 5% level of confidence. Across every specification, the positive coefficient on the interaction term *CASH_FLOW*×*FUNCTIONAL_DISTANCE* indicates that the marginal effect of cash flows on investment is increasing with functional distance. It is worth noting that even if the coefficient on cash flow is negative (albeit insignificant), the estimated marginal effect of *CASH_FLOW* is positive and the corresponding elasticity of investment to cash flow for the average firm is, as expected, around unity.[23]

[22] Since we have data on *RAT* only for the year of the survey, we assume that the same rationing condition holds for the entire three-year period and calculate robust standard error. Results are available in the working paper version of this paper (Alessandrini et al., 2006).

[23] From equation (2), it is easy to calculate the threshold value of *FUNCTIONAL_DISTANCE* indexes beyond which the marginal effect of cash flows on investment becomes positive: the corresponding values for *FUNCTIONAL_DISTANCE_KM* and *FUNCTIONAL_DISTANCE_LATENT* are below the second percentile, showing that the marginal effect $\partial(INVESTMENT/K)/\partial(CASH\_FLOW/K)$ is positive in at least 98% of cases.

Table VIII. Investment-cash flow sensitivity: One Step System-GMM estimation of Equation (2)

The table reports regression coefficients and, in brackets, the associated standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable *(INVESTMENT/K)* is the ratio between investment and the capital stock. *GROWTH* is the annual growth rate (in percentage) of total sales, *(CASH_FLOW/K)* is the ratio between cash flow and capital stock. Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j* (*FUNCTIONAL_DISTANCE_KM*) and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches (× 10,000) over surface (*OPERATIONAL_PROXIMITY_AREA*). *SIZE* is the number of workers categorized if five classes: 11–20 workers ($SIZE_1$, taken as benchmark), 21–50 workers ($SIZE_2$), 51–250 workers ($SIZE_3$), 251–499 workers ($SIZE_4$) and more than 500 workers ($SIZE_5$). *(CASH_FLOW/K)×FUNCTIONAL_DISTANCE×*$SIZE_i$ refers to the interaction with *FUNCTIONAL_DISTANCE_KM* (columns 5–7) and *FUNCTIONAL_DISTANCE_LATENT* (columns 6–8). All regressions include the constant and time, size, geographic (five macro-areas) and industry (Pavitt classification) dummies. The model is estimated by One-Step System GMM using Stata 10 SE package. As instruments, we use lagged values of endogenous variables from t-1 to t-5. Geographic, industry and time dummies are taken as strictly exogenous regressors. As diagnostic, the table reports the p-values of the Wald-test for the significance of the regression, of the Hansen test for over-identifying restrictions (OIR, the null is the validity of the instrument set) and of the Arellano and Bond autocorrelation tests of first and second order (AR(1) and AR(2), the null is no autocorrelation). The bottom lines report the estimated elasticity of investment to cash flow: the first four column refer to the average firm' elasticity, while the last four columns report the elasticities according to the firm dimensional class, reckoned at the average values of *(INVESTMENT/K)*, *(CASH_FLOW/K)*, *FUNCTIONAL_DISTANCE_KM*, *FUNCTIONAL_DISTANCE_LATENT*, *OPERATIONAL_PROXIMITY_POPULATION* and *OPERATIONAL_PROXIMITY_AREA*. The asterisks report the significance level of a two tailed t-test of annulment.

Dep Var: (*INVESTMENT/K*)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(*INVESTMENT/K*)(-1)	0.002***	0.002***	0.002***	0.002***	0.002***	0.002**	0.002**	0.002**
	[0.000]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]	[0.001]	[0.001]
GROWTH(-1)	0.147***	0.147***	0.150***	0.149***	0.152***	0.150***	0.154***	0.154***
	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]	[0.008]	[0.008]
(*CASH_FLOW/K*)	−2.086**	−2.030**	0.315	0.292	−2.190***	−1.982***	−0.718	−0.595
	[0.856]	[0.866]	[0.244]	[0.220]	[0.753]	[0.742]	[0.820]	[0.669]
(*CASH_FLOW/K*) × *FUNCTIONAL_DISTANCE_KM*	0.456***		0.177*		0.816***		0.635*	
	[0.167]		[0.095]		[0.186]		[0.352]	

Table VIII. (Continued)

Dep Var: (INVESTMENT/K)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(*CASH_FLOW/K*) × *FUNCTIONAL_DISTANCE_LATENT*		0.575***		0.253**		0.960***		0.802**
		[0.200]		[0.114]		[0.219]		[0.392]
(*CASH_FLOW/K*) × *OPERATIONAL_PROXIMITY_POPULATION*	0.245**	0.252**			0.167*	0.177*		
	[0.113]	[0.114]			[0.098]	[0.095]		
(*CASH_FLOW/K*) × *OPERATIONAL_PROXIMITY_AREA*			−0.004**	−0.004**			−0.002	−0.002
			[0.001]	[0.001]			[0.004]	[0.003]
(*CASH_FLOW/K*) × *FUNCTIONAL_DISTANCE*×$SIZE_2$					−0.111**	−0.132*	−0.063	−0.079
					[0.054]	[0.077]	[0.121]	[0.171]
(*CASH_FLOW/K*) × *FUNCTIONAL_DISTANCE*×$SIZE_3$					−0.195**	−0.227*	−0.085	−0.100
					[0.095]	[0.119]	[0.106]	[0.139]
(*CASH_FLOW/K*) × *FUNCTIONAL_DISTANCE*×$SIZE_4$					−0.160**	−0.223*	−0.112	−0.141
					[0.078]	[0.133]	[0.069]	[0.114]
(*CASH_FLOW/K*) × *FUNCTIONAL_DISTANCE*×$SIZE_5$					−0.412***	−0.677***	−0.425**	−0.691***
					[0.142]	[0.179]	[0.192]	[0.266]
Observations	1,953	1,953	1,953	1,953	1,953	1,953	1,953	1,953
Number of firms	279	279	279	279	279	279	279	279
OIR test	0.184	0.234	0.178	0.112	0.289	0.313	0.122	0.167
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.639	0.627	0.641	0.636	0.651	0.629	0.669	0.650
Wald-test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Elasticities of investment to cash flow according to firm size:								
Average firm	1.014***	1.047***	1.117***	1.133***				
11–20 workers					1.198***	1.132***	1.127***	1.101***
21–50 workers					1.413***	1.450***	1.469***	1.532***
51–250 workers					0.845**	0.918**	1.209***	1.266***
251–499 workers					0.842***	0.817**	0.863**	0.913**
More than 500 workers					−0.239	−0.591	−0.308	−0.612

With respect to the operational proximity indexes, we find that *OPERATIONAL_PROXIMITY_POPULATION* and *OPERATIONAL_PROXIMITY_AREA* have opposite effects on investment-cash flow sensitivity, with the former raising and the latter easing the sensitivity of investment to cash flow, even if this effect vanishes in the augmented specification. Altogether, these results confirm the results of previous literature: namely, the ambiguous effects that the number of banks operating in a region may have on local borrowers.

The results of the augmented specification point to significant differences in the impact of functional distance on investment-cash flow sensitivity according to firm size: from columns 5 and 6, we can observe that the larger the firm, the lower is the contribution of functional distance to the marginal effect of cash flow on investment. This relation is non-linear and is especially weak for the largest firms. When considering *FUNCTIONAL_DISTANCE_LATENT* (columns 7 and 8), the impact on $\partial(INVESTMENT/K)/\partial(CASH\_FLOW/K)$ is significantly lower only for firms employing more than 500 workers. The diminishing marginal effect of functional distance is reflected in a declining path of the elasticity of investment to cash flow, as shown by the last rows of Table VIII. In particular, investment choices are very elastic to internal financing for firms employing less than 50 workers, while the relationship becomes inelastic for large firms with more than 500 workers.

6. Credit lines drawn and overdrawn

6.1 THE ECONOMETRIC MODEL

The last two empirical exercises are carried out at the market level and consider as dependent variables the percentage of credit drawn from available credit lines (*CREDIT_DRAWN*) and the ratio of overdrawn to actual credit granted (*OVERDRAWN*). The former is a measure of liquidity constraint, while the latter is an indicator of the flexibility of credit supply, i.e., the ability (capacity) of the banking system to provide borrowers with liquidity. Therefore, in a functionally distant banking system that tightens local credit market conditions, functional distance indexes must be positively correlated with *CREDIT_DRAWN* and negatively correlated with *OVERDRAWN*.

Data on *CREDIT_DRAWN* and *OVERDRAWN* are available at the provincial level by borrower location for the period 1997–2003. Given the high time persistence shown by these indexes, we estimate a dynamic balanced panel model, in which the dependent variable (*CREDIT_DRAWN* and *OVERDRAWN*, alternatively) is regressed against its lagged value, functional distance and operational proximity indexes, and n control variables at the provincial level (*PROV*):

$$\ln Y_{jt} = \alpha + \beta_1 \ln Y_{jt-1} + \beta_2 OPERATIONAL\_PROXIMITY_{jt} + \beta_3 FUNCTIONAL\_DISTANCE_{jt} + \sum_n \varphi_n PROV_{njt} + \eta_j + \nu_t + \varepsilon_{jt} \quad (4)$$

To be consistent with previous exercises, we include in *PROV* control variables for credit market concentration (*HHI*) and localism (*COOPERATIVE_BANKS*). As a measure of local economic development, we add the logarithm of real per capita value added (*VALUE_ADDED*) to capture the average degree of creditworthiness of local borrowers, and we include the dummies for the five Italian macro-regions. Finally, η_j and ν_t take into account time and individual fixed effects, whereas ε_{jt} is the idiosyncratic error term.

Given the dynamic structure of the model, and to control also for potential endogeneity of right-hand side variables, we estimate Equation (4) using the System-GMM. We consider separate models for three loan size classes, as a proxy for borrowers' size in order to test whether distances matter more for small than for large firms.[24]

6.2 RESULTS

6.2.1 Credit Lines Drawn

Consistent with previous analysis, estimation results of model (4) provides evidence of a positive and significant relationship between functional distance and credit shortage limited to small and medium loan classes (Table IX). All specifications pass the standard Hansen test of overidentification, suggesting that the lag structure of the instruments is correct. Moreover, we reject the hypothesis of no first-order autocorrelation, while we cannot reject that of no second-order autocorrelation.

Considering distance variables, functional distance makes the local credit market tighter for small borrowers, regardless of the *FUNCTIONAL_DISTANCE* indexes used. The steady-state elasticity of *CREDIT_DRAWN* with respect to functional distance ranges from 0.06 to 0.11. By contrast, we find that operational proximity has ambiguous effects on *CREDIT_DRAWN*. Relying on branches per capita, we find that the higher the *OPERATIONAL_PROXIMITY_POPULATION*, the lower is the share of available credit lines (the elasticity being −0.09, as from column 1 of Table IX; result not reported). By contrast, the coefficient of *OPERATIONAL_PROXIMITY_AREA* is positive and statistically significant, even if its

[24] Loan classes are: € 75,000–125,000; € 125,000–500,000; more than € 500,000. For robustness we also consider a five loan class classification where loans above € 125,000 are divided into € 125,000–250,000; € 250,000–500,000; € 500,000–2,500,000; more than € 2,500,000.

Table IX. Credit lines drawn: One Step System-GMM estimation of Equation (4)

The table reports regression coefficients and, in brackets, the associated robust standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable (*CREDIT_DRAWN*) is the logarithm of the ratio between utilized and available credit lines. *CREDIT_DRAWN*(-1) is the lagged value of the dependent variable. To obtain more readable coefficients, *CREDIT_DRAWN* and *CREDIT_DRAWN*(-1) are multiplied by 10. Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j* (*FUNCTIONAL_DISTANCE_KM*) and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches ($\times$ 10,000) over surface (*OPERATIONAL_PROXIMITY_AREA*). *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province. *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province. *VALUE_ADDED* is the logarithm of per capita real value added in the province. All regressions include the constant and time and geographic (five macro-areas) dummies. The model is estimated by One-Step System GMM using Stata 10 SE package. As instruments, we use lagged values (t-1 and t-2) of *CREDIT_DRAWN*(-1) and the available lagged values after t-2 for the other variables. Geographic and time dummies are taken as strictly exogenous regressors. As diagnostic, the table reports the p-values of the Wald-test for the significance of the regression, of the Hansen test for over-identifying restrictions (OIR, the null is the validity of the instrument set) and of the Arellano and Bond autocorrelation tests of first and second order (AR(1) and AR(2), the null is no autocorrelation). The bottom lines report the estimated long-run elasticity of *CREDIT_DRAWN* with respect to *FUNCTIONAL_DISTANCE* indexes, reckoned at the average values of *FUNCTIONAL_DISTANCE_KM* and *FUNCTIONAL_DISTANCE_LATENT*.

Dep Var: *CREDIT_DRAWN*	Loan size (thousands €)			Loan size (thousands €)			Loan size (thousands €)			Loan size (thousands €)		
	75–125	125–500	>500	75–125	125–500	>500	75–125	125–500	>500	75–125	125–500	>500
CREDIT_DRAWN(-1)	0.613***	0.827***	0.286**	0.617***	0.828***	0.286**	0.614***	0.824***	0.284**	0.623***	0.826***	0.286**
	[0.092]	[0.044]	[0.140]	[0.091]	[0.045]	[0.138]	[0.094]	[0.042]	[0.139]	[0.092]	[0.043]	[0.138]
FUNCTIONAL_DISTANCE_KM	0.097**	0.057**	−0.036				0.085**	0.045	−0.029			
	[0.042]	[0.027]	[0.158]				[0.039]	[0.028]	[0.161]			

Table IX. (Continued)

Dep Var: (*INVESTMENT/K*)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)				
FUNCTIONAL_DISTANCE_ LATENT				0.121**	0.070**	0.033				0.085**	0.063*	0.061
				[0.056]	[0.033]	[0.191]				[0.043]	[0.037]	[0.188]
OPERATIONAL_PROXIMITY_ POPULATION	−0.063**	−0.021	−0.201	−0.058	−0.020	−0.197						
	[0.032]	[0.026]	[0.158]	[0.036]	[0.027]	[0.155]						
OPERATIONAL_PROXIMITY_ AREA							0.003**	0.002***	−0.001	0.003**	0.002**	−0.001
							[0.001]	[0.001]	[0.004]	[0.001]	[0.001]	[0.004]
COOPERATIVE_BANKS	0.242	0.409	0.623	0.321	0.417	0.985	−0.066	0.348	−0.485	−0.174	0.370	0.000
	[0.392]	[0.282]	[1.330]	[0.439]	[0.298]	[1.343]	[0.368]	[0.359]	[1.541]	[0.390]	[0.397]	[1.494]
HHI	0.687**	0.817***	2.454***	0.674*	0.791***	2.539***	0.561*	0.853***	2.009**	0.533	0.848***	2.107**
	[0.335]	[0.244]	[0.875]	[0.358]	[0.255]	[0.904]	[0.312]	[0.267]	[0.862]	[0.326]	[0.289]	[0.931]
VALUE_ADDED	0.559**	0.137	0.263	0.559*	0.136	0.326	0.157	−0.082	−0.35	0.131	−0.064	−0.227
	[0.254]	[0.153]	[0.841]	[0.288]	[0.162]	[0.815]	[0.211]	[0.180]	[1.001]	[0.218]	[0.188]	[0.939]
Observations	618	618	618	618	618	618	618	618	618	618	618	618
OIR test	0.454	0.337	0.310	0.665	0.354	0.399	0.397	0.319	0.210	0.484	0.309	0.224
AR(1)	0.108	0.002	0.021	0.108	0.002	0.020	0.108	0.001	0.024	0.107	0.001	0.023
AR(2)	0.166	0.789	0.197	0.167	0.788	0.188	0.168	0.781	0.149	0.165	0.778	0.142
Wald-test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Elasticity wrt *FUNCTIONAL_ DISTANCE*	0.082	0.107	−0.017	0.081	0.105	0.012	0.072	0.083	−0.013	0.058	0.093	0.022

economic effect is negligible (the elasticity is 0.01, as from column 7 of Table IX; result not reported).

Moving on to the control variables, we find strong evidence of path dependence for *CREDIT_DRAWN*, while, contrary to what we expected, the coefficient of per capita value added is either significantly positive or insignificant. As regards the characteristics of the local credit market, we find that the *HHI* coefficient is positive and significant, while the share of branches held by credit cooperative banks in the provinces does not have any effect on *CREDIT_DRAWN*.

6.2.2 Overdrawn

Table X reports the results of model (4) with the overdrawn ratio as a dependent variable. The diagnostic tests support the model specification and the lag structure of the instruments. Focusing on distance variables, branch density has no effect on *OVERDRAWN*, while functional distance, however measured, has a significant negative impact for the smallest loan class. In particular, the high elasticity of *OVERDRAWN* to functional distance, which ranges between 0.42 and 0.55, suggests that the proximity of the banks' decisional centres has a major role in determining the flexibility of credit supply for small borrowers.

The coefficients on *VALUE_ADDED* and *HHI* are consistent with previous results on credit lines drawn. In richer provinces and in credit markets where the banking system is more concentrated, credit supply is less flexible to the immediate liquidity needs of small borrowers who show restrained opportunities to overdraw. Furthermore, we also find that, at the aggregate level, a larger share of credit cooperative banks are associated with a tighter credit market.

7. Robustness Checks

We undertake a number of robustness checks concerning the functional distance indexes, econometric specification, and the estimation methodology.[25] Overall, results on the distance variables appear to be robust across all these additional exercises.

With regard to functional distance, assuming that in the case of multibank holdings power on lending policy is in the hands of the parent bank, we use *FUNCTIONAL_DISTANCE_G* indexes based on distance between the local branch and the headquarter of the parent bank. In all exercises the results are qualitatively similar to those obtained with *FUNCTIONAL_DISTANCE* indexes.

With regard to econometric specification, in the credit rationing model we estimate Equation (1) with contemporaneous or one year lagged firm-specific variables

[25] To save space, results and tables are not reported but are available on request from the authors.

Table X. Overdraft: One Step System-GMM estimation of Equation (4)

The table reports regression coefficients and, in brackets, the associated robust standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable (*OVERDRAWN*) is the logarithm of the ratio between utilized and available credit lines. *OVERDRAWN*(-1) is the lagged value of the dependent variable. To obtain more readable coefficients, *OVERDRAWN* and *OVERDRAWN*(-1) are multiplied by 10. Functional distance indexes are computed as the ratio of branches weighted by the logarithm of 1 plus the kilometric distance between the province of the branch and that where the parent bank is headquartered, over total branches in province *j* (*FUNCTIONAL_DISTANCE_KM*) and as a latent distance (*FUNCTIONAL_DISTANCE_LATENT*) obtained by taking the first component from a Factorial Matrices Analysis run on the three basic functional distance indexes. The operational proximity indexes are computed as the number of bank branches in the province per 10,000 inhabitants (*OPERATIONAL_PROXIMITY_POPULATION*), and the ratio of branches (× 10,000) over surface (*OPERATIONAL_PROXIMITY_AREA*). *COOPERATIVE_BANKS* is the share of branches held by Credit Cooperative Banks on total branches in the province. *HHI* is the logarithm of the Herfindahl-Hirschman Index calculated on the number of branches in the province. *VALUE_ADDED* is the logarithm of per capita real value added in the province. All regressions include the constant and time and geographic (five macro-areas) dummies. The model is estimated by One-Step System GMM using Stata 10 SE package. As instruments, we use all the available lagged values after t-3 of the endogenous variables and the lagged values in t-1, t-2 and t-3 of in the *FUNCTIONAL_DISTANCE* indexes. Geographic and time dummies are taken as strictly exogenous regressors. As diagnostic, the table reports the p-values of the Wald-test for the significance of the regression, of the Hansen test for over-identifying restrictions (OIR, the null is the validity of the instrument set) and of the Arellano and Bond autocorrelation tests of first and second order (AR(1) and AR(2), the null is no autocorrelation). The bottom lines report the estimated long-run elasticity of *OVERDRAWN* with respect to *FUNCTIONAL_DISTANCE* indexes, reckoned at the average values of *FUNCTIONAL_DISTANCE_KM* and *FUNCTIONAL_DISTANCE_LATENT*.

Dep Var: *OVERDRAWN*	Loan size (thousands €)			Loan size (thousands €)			Loan size (thousands €)			Loan size (thousands €)		
	75–125	125–500	>500	75–125	125–500	>500	75–125	125–500	>500	75–125	125–500	>500
OVERDRAWN(-1)	0.564***	0.661***	0.392**	0.572***	0.660***	0.397***	0.531***	0.647***	0.417**	0.529***	0.645***	0.430**
	[0.086]	[0.073]	[0.154]	[0.086]	[0.071]	[0.151]	[0.092]	[0.073]	[0.170]	[0.093]	[0.074]	[0.168]
FUNCTIONAL_DISTANCE_KM	−0.731**	0.075	0.733				−0.605*	0.198	0.476			
	[0.347]	[0.292]	[0.493]				[0.337]	[0.321]	[0.513]			
FUNCTIONAL_DISTANCE_LATENT				−0.876**	0.209	0.982				−0.769*	0.280	0.648
				[0.415]	[0.387]	[0.614]				[0.442]	[0.432]	[0.675]

Table X. (Continued)

Dep Var: *OVERDRAWN*	Loan size (thousands €)			Loan size (thousands €)			Loan size (thousands €)			Loan size (thousands €)		
	75–125	125–500	>500	75–125	125–500	>500	75–125	125–500	>500	75–125	125–500	>500
OPERATIONAL_PROXIMITY_POPULATION	0.224	−0.140	0.112	0.265	−0.098	0.102						
	[0.239]	[0.205]	[0.444]	[0.257]	[0.205]	[0.464]						
OPERATIONAL_PROXIMITY_AREA							−0.005	0.012	0.034*	−0.006	0.014	0.038*
							[0.012]	[0.013]	[0.020]	[0.012]	[0.014]	[0.021]
COOPERATIVE_BANKS	−9.136***	−4.305	−1.614	−8.828***	−3.661	−0.727	−9.318**	−5.402	−4.660	−9.311**	−5.054	−3.608
	[3.186]	[3.016]	[3.813]	[3.255]	[2.992]	[3.894]	[3.642]	[3.474]	[4.731]	[3.966]	[3.703]	[5.095]
HHI	−3.519**	−0.937	1.005	−3.396**	−0.738	1.109	−3.798**	−1.732	1.187	−3.837**	−1.484	1.541
	[1.712]	[1.671]	[2.167]	[1.687]	[1.709]	[2.147]	[1.681]	[1.764]	[2.126]	[1.679]	[1.832]	[2.119]
VALUE_ADDED	−5.654***	−4.046***	−6.104*	−5.858***	−4.350***	−6.142*	−5.306**	−5.283***	−7.294*	−5.479**	−5.497***	−7.488*
	[1.801]	[1.469]	[3.480]	[1.847]	[1.482]	[3.532]	[2.398]	[1.827]	[3.938]	[2.535]	[1.826]	[3.928]
Observations	618	618	618	618	618	618	618	618	618	618	618	618
OIR test	0.471	0.260	0.289	0.485	0.299	0.387	0.498	0.354	0.563	0.516	0.422	0.617
AR(1)	0.001	0.002	0.001	0.001	0.002	0.001	0.001	0.004	0.001	0.001	0.004	0.001
AR(2)	0.086	0.265	0.311	0.085	0.270	0.308	0.084	0.280	0.297	0.086	0.282	0.288
Wald-test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Elasticity wrt *FUNCTIONAL_DISTANCE*	−0.547	0.073	0.393	−0.524	0.158	0.417	−0.421	0.183	0.266	−0.418	0.202	0.291

and include additional controls like *ROE* (whose coefficient is significantly positive), firm age, length of the bank-firm relationship and the quadratic terms of the Herfindahl-Hirschman index (not significant). None of these changes modify the sign or significance of distance coefficients. In the investment-cash flow sensitivity analysis we find that the results are unaffected by the inclusion of the firm's leverage (*DEBT*), which has a positive but not significant effect on investment, and by geographical dummies at the regional level. In the credit lines drawn and overdrawn regressions we include *FAIL* as an additional control variable which is not significant and does not affect the positive correlation between functional distance and credit tightness. Moreover, we control for geographic fixed effects at the regional level.

Finally, with regard to econometric methodology, in the probit analysis we try to exploit the time dimension of the dataset, building a balanced panel for 195 firms over three time periods, corresponding to the years of the surveys. Random effect logit estimation confirms the negative sign of the coefficient of functional distance, even if it is no longer significant. As already noted, this result might be due to the reduced sample size and to the very limited number of changes in rationed status by firms included in our sample. Besides, the rotating panel of Capitalia survey suffers from a high degree of attrition that creates a bias in the estimates, possibly outweighing the efficiency gains of exploiting unobserved firm-specific heterogeneity (Nese and O'Higgins 2007). Finally, we are further persuaded about the validity of the pooled estimator by the fact that we cannot always reject at 5% level of significance the hypothesis that the fraction of total variance contributed by the panel-level variance component, the rho index, is equal to zero.

We estimate the investment equation also with a static specification with the within-group estimator. Results confirm the positive effect of cash flow on investment and also the positive marginal effect of functional distance on the sensitivity of investment to cash flow. However, controlling for firm-specific invariant effects, we find that operational proximity reduces the marginal impact of cash flow on investment. Finally, both in the investment-cash flow, and credit drawn and overdrawn exercises our findings are generally confirmed by using a two-step system-GMM estimator.

8. Conclusions

Bank deregulations, progress in information technology and incessant financial innovation have stimulated geographical diffusion of banking structures and instruments that have greatly increased operational proximity of banks to local economies. Working in the opposite direction, the consolidation of the banking industry has also produced a geographical concentration of decision and strategic centres of banking institutions that has tended to increase functional distance of the banking system from local communities.

In this paper, we assess the impact of this spatial diffusion-concentration trends on financing constraints for firms, focusing on the Italian experience over the period 1996–2003. Our analysis consistently shows that greater functional distance has hardened firms' financing constraints. Functional distance is positively associated with the probability of firms being rationed, investment-cash flow sensitivity and share of credit lines actually drawn, and negatively associated with overdrawing. These adverse effects are particularly evident for small firms and for firms located in less developed southern Italian provinces.

Although one should be wary of extending country-specific results to other contexts, our findings suggest a more general policy conclusion. The intense bank consolidation process in the last twenty years has not made retail credit markets fully integrated even at the country level. On the contrary, it has made access to credit for informationally opaque firms in peripheral regions still harder. As Mortimer-Schutts (2005) has recently pointed out, in order to bring the benefits of market deregulation to local borrowers, what is needed is to promote the entry and creation of new banks and non-bank competitors struggling to capture a share of the local credit market by offering new services and not by merely incorporating local incumbents. Only by pursuing this track firmly will it be possible for banking competition to focus on the real needs of local economic development.

Funding

Ministero dell'Università e della Ricerca, "The new geography of the Italian financial system: Markets, intermediaries and development" (Project Nr. 2006130584).

References

Agarwal, S., Ambrose, B.W. and Liu, C. (2006) Credit lines and credit utilization, *Journal of Money Credit and Banking* **38**, 1–22.

Agarwal, S. and Hauswald, R. (2008) Distance and information asymmetries in lending, unpublished working paper, Federal Reserve Bank of Chicago, American University.

Aghion, P. and Tirole, J. (1997) Formal and real authority in organizations, *Journal of Political Economy* **105**, 1–29.

Alessandrini, P., Calcagnini, G. and Zazzaro, A. (2008) Asset restructuring strategies in bank acquisitions: Does distance between dealing partners matter?, *Journal of Banking and Finance*, forthcoming.

Alessandrini, P., Croci, M. and Zazzaro, A. (2005) The geography of banking power: the role of functional distance, *BNL Quarterly Review* **235**, 129–167.

Alessandrini, P., Presbitero, A.F. and Zazzaro, A. (2006) Banks, distances and financing constraints for firms, unpublished working paper, Università Politecnica delle Marche.

Alessandrini, P. and Zazzaro, A. (1999) A "possibilist" approach to local financial systems and regional development: The Italian experience, in: Martin R. (ed.), *Money and the Space Economy*. Wiley & Sons, New York, pp. 71–92.

Angelini, P., Di Salvo, R. and Ferri, G. (1998) Availability and cost of credit for small businesses: Customer relationships and credit cooperatives, *Journal of Banking and Finance* **22**, 925–954.

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

Avery, R.B. and Samolyk, K. (2004) Bank consolidation and the provision of banking services: Small commercial loans, *Journal of Financial Services Research* **25**, 291–325.

Benfratello, L., Schiantarelli, F. and Sembenelli, A. (2008) Banks and innovation: microeconometric evidence on Italian firms, *Journal of Financial Economics*, forthcoming.

Berger, A.N., Demsetz, R.S. and Strahan, P.E. (1999) The consolidation of financial services industry: Causes, consequences, and implications for the future, *Journal of Banking and Finance* **23**, 135–194.

Berger, A.N. and DeYoung, R. (2006) Technological progress and the geographic expansion of the banking industry, *Journal of Money Credit and Banking* **38**, 1483–1514.

Berger, A.N., Miller, N.H., Petersen, M.A., Rajan, R.G. and Stein, J.C. (2005) Does function follow organizational form? Evidence from the lending practices of large and small banks, *Journal of Financial Economics* **76**, 237–269.

Berger, A.N., Rosen, R.J. and Udell, G.F. (2007) Does market size structure affect competition? The case of small business lending, *Journal of Banking and Finance* **31**, 11–33.

Berger, A.N., Saunders, A., Scalise, J.M. and Udell, G.F. (1998) The effects of bank mergers and acquisitions on small business lending, *Journal of Financial Economics* **50**, 187–229.

Berger, A.N. and Udell, G.F. (1996) Universal banking and the future of small business lending, in: A. Saunders and I. Walter (eds.), *Universal Banking*. Financial System Design Reconsidered. Irwin, Chicago, pp- 559–627.

Berger, A.N. and Udell, G.F. (2002) Small business credit availability and relationship lending: The importance of bank organizational structure, *Economic Journal* **112**, F32–F53.

Berger, A.N. and Udell, G.F. (2006) A more complete conceptual framework for SME finance, *Journal of Banking and Finance* **30**, 2945–2966.

Blundell, R. and Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics* **87**, 115–143.

Bonaccorsi di Patti, E. (2003) Preferences in bank-firm relationships, unpublished working paper, Bank of Italy.

Bonaccorsi di Patti, E. and Dell'Ariccia, G. (2004) Bank competition and firm creation, *Journal of Money Credit and Banking* **36**, 225–251

Bonaccorsi di Patti, E. and Gobbi, G. (2001) The changing structure of local credit markets: Are small business special?, *Journal of Banking and Finance* **25**, 2209–2237.

Bonaccorsi di Patti, E. and Gobbi, G. (2007) Winners or losers? The effects of banking consolidation on corporate borrowers, *Journal of Finance* **62**, 669–695.

Boot, A.W.A. and Thakor, A.V. (2000) Can relationship banking survive competition?, *Journal of Finance* **55**, 679–713.

Brevoort, K.P. and Hannan, T.H. (2006) Commercial lending and distance: evidence from Community Reinvestment Act data, *Journal of Money Credit and Banking* **38**, 1991–2012.

Brickley, J.A., Linck, J.S. and Smith, C.W. (2003) Boundaries of the firm: evidence from the banking industry, *Journal of Financial Economics* **70**, 351–383.

Broecker, T. (1990) Credit-worthiness tests and interbank competition, *Econometrica* **58**, 429–452.

Capitalia (2005). Indagine sulle imprese manifatturiere. *Nono Rapporto sull'industria italiana e sulla politica industriale*, Capitalia, Rome.
Carling, K. and Lundberg, S. (2005) Asymmetric information and distance: an empirical assessment of geographical credit rationing, *Journal of Economics and Business* **57**, 39–59.
Casolaro, L. and Mistrulli, P.E. (2007) Distance, organizational structure and loan rates, unpublished working paper, Bank of Italy.
Cole, R., Goldberg, L. and White, L. (2004) Cookie-cutter versus character: The micro structure of small business lending by large and small banks, *Journal of Financial and Quantitative Analysis* **39**, 227–251.
Collender, R.N. and Shaffer, S. (2003) Local bank ownership, deposit control, market structure, and economic growth, *Journal of Banking and Finance* **27**, 27–57.
Degryse, H. and Ongena, S. (2001) Bank relationships and firm profitability, *Financial management* **30**, 9–34.
Degryse, H. and Ongena, S. (2005) Distance, lending relationship, and competition, *Journal of Finance* **60**, 231–266.
Dell'Ariccia, G. and Marquez, R. (2004) Information and bank credit allocation, *Journal of Financial Economics* **72**, 185–214.
Dessein, W. (2002) Authority and communications in organizations, *Review of Economic Studies* **69**, 811–838.
Fazzari, S.M., Hubbard, R.G. and Petersen, B.C. (1988) Financing constraints and corporate investments, *Brooking Papers on Economic Activity* **1**, 141–195.
Fachin, S., Shea, D.G. and Vichi, M. (2002) Exploring 3D datasets: A factorial matrices analysis of the US industry in the 1980s, *Applied Economics* **34**, 295–304.
Gehrig, T. (1998) Screening, cross-border banking, and the allocation of credit, *Research in Economics* **52**, 387–407.
Guelpa, F. and Tirri, V. (2006) The effect of market structure and relationship lending on the likelihood of credit tightening, unpublished working paper, Banca Intesa.
Guiso, L., Sapienza, P. and Zingales, L. (2004a) Does local financial development matter?, *The Quarterly Journal of Economics* **119**, 929–969.
Guiso, L., Sapienza, P. and Zingales, L. (2004b) The role of social capital in financial development, *American Economic Review* **94**, 526–556
Harris, M. and Raviv, A. (2005) Allocation of decision-making authority, *Review of Finance* **9**, 353–383.
Hauswald, R.B.H. and Marquez, R. (2003) Information technology and financial services competition, *The Review of Financial Studies* **16**, 921–948.
Hauswald, R.B.H. and Marquez, R. (2006) Competition and strategic information acquisition in credit markets, *The Review of Financial Studies* **19**, 967–1000.
Herrera, A.M. and Minetti, R. (2007) Informed finance and technological change: Evidence from credit relationships, *Journal of Financial Economics* **83**, 223–269.
Hertzberg, A., Liberti, J.M. and Paravisini, D. (2007) Information and incentives inside the firm: Evidence from loan officer rotation, unpublished working paper, Columbia Business School.
Hirshleifer, D.A and Thakor, A.V. (1992) Managerial conservatism, project choice, and debt, *Review of Financial Studies* **5**, 437–470.
Huynh, K.P. and Rotondi, Z. (2007) R&D spending, local banks, and knowledge spillovers, unpublished working paper, Indiana University, Ferrara University.
Ichino, A. and Maggi, G. (2000) Work environment and individual background: Explaining regional shirking differential in a large Italian firm, *The Quarterly Journal of Economics* **115**, 1057–1090.

Inderst, R. and Mueller, H.M. (2007) A lender-based theory of collateral, *Journal of Financial Economics* **84**, 826–859.

ISTAT (2001) *8° Censimento generale dell'industria e dei servizi 2001*, ISTAT, Rome.

Jenness, J. (2005) Distance matrix (dist_mat_jen.avx) extension for ArcView 3.x, v. 2, *Jenness Enterprises*, available at: http://www.jennessent.com/arcview/dist_matrix.htm.

Jimenez, G., Salas, V. and Saurina, J. (2007) Organizational distance and use of collateral in business loans, unpublished working paper, Bank of Spain.

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

Keeton, W.R. (1995) Multi-office bank lending to small businesses: Some new evidence, *Federal Reserve Bank of Kansas City Economic Review*, Second Quarter, 45–57.

Liberti, J.M. (2004) Initiative, incentives and soft information. How does delegation impact the role of bank relationship managers?, unpublished working paper, London Business School.

Liberti, J.M. and Mian, A. (2006) Estimating the effect of hierarchies on information use, unpublished working paper, London Business School.

Manove, M., Padilla, A.J. and Pagano, M. (2001) Collateral versus project screening: A model of lazy banks, *RAND Journal of Economics* **32**, 726–744.

McAfee, R.P. and McMillan, J. (1995) Organizational diseconomies of scale, *Journal of Economics & Management Strategy* **4**, 399–426.

Mian, A. (2006) Distance constraints: The limits of foreign lending in poor economies, *Journal of Finance* **61**, 1465–1505.

Milgrom, P. and Roberts, J. (1990) Bargaining costs, influence costs, and the organization of economic activity, in: Alt J. and Shepsle K. (eds.), *Perspective on Positive Political Economy*. Cambridge University Press, Cambridge, pp. 57–89.

Montoriol Garriga, J. (2006) The effect of relationship lending on firm performance, unpublished working paper, Universitat Autònoma de Barcelona.

Mortimer-Schutts, I. (2005) EU regulatory and supervisory convergence: The case for a dual system with choice, Policy Discussion Paper, AEI-Brookings Joint Center for Regulatory Studies.

Nese, A. and O'Higgins, N. (2007) Attrition bias in the Capitalia panel, *International Review of Economics* **54**, 383–403.

Nickell, S. (1981) Biases in dynamic models with fixed effects, *Econometrica* **49**, 1417–1426.

Novaes, W. and Zingales, L. (2004) Bureaucracy as a mechanism to generate information, *RAND Journal of Economics* **35**, 245–259.

Ongena, S. and Smith, D.C. (2000) Bank relationships: A review, in: Zenios S.A. and Harker P. (eds.), *Performance of Financial Institutions*. Cambridge University Press, Cambridge, pp. 221–258.

Palley, T.L. (1997) Managerial turnover and the theory of short-termism, *Journal of Economic Behavior and Organization* **32**, 547–557.

Peek, J. and Rosengren, E.S. (1998) Bank consolidation and small business lending: It's not just bank size that matters, *Journal of Banking and Finance* **22**, 799–819.

Petersen, M.A. and Rajan, R.G. (1994) The benefits of lending relationships: Evidence from small business data, *The Journal of Finance* **49**, 3–37.

Petersen, M.A. and Rajan, R.G. (1995) The effect of credit market competition on lending relationships, *The Quarterly Journal of Economics* **110**, 407–443.

Petersen, M.A. and Rajan, R.G. (2002) Does distance still matter? The information revolution in small business lending, *The Journal of Finance* **57**, 2533–2570.

Sapienza, P. (2002) The effects of banking mergers on loan contracts, *The Journal of Finance* **57**, 329–367.

Stein, J.C. (2002) Information production and capital allocation: Decentralized versus hierarchical firms, *The Journal of Finance* **57**, 1891–1921.

Scharfstein, D.S. and Stein, J.C. (2000) The dark side of internal capital markets: Divisional rent-seeking and inefficient investment, *The Journal of Finance* **55**, 2537–2564.

Shaffer, S. (1998) The winner's curse in banking, *Journal of Financial Intermediation* **7**, 359–392.

Takáts, E. (2004) Banking consolidation and small business lending, European Central Bank Working Paper No. 407.

Udell, G.F. (1989) Loan quality, commercial loan review and loan officer contracting, *Journal of Banking and Finance* **13**, 367–382.

Zazzaro, A. (2005) Should courts enforce credit contracts strictly?, *Economic Journal* **115**, 166–184.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Banking & Finance 30 (2006) 2945–2966
Journal of BANKING & FINANCE
www.elsevier.com/locate/jbf

A more complete conceptual framework for SME finance ☆

Allen N. Berger [a,b,*], Gregory F. Udell [c]

[a] *Board of Governors of the Federal Reserve System, Mail Stop 153, Federal Reserve Board, 20th and C Streets, Washington, DC 20551, USA*
[b] *Wharton Financial Institutions Center, Philadelphia, PA 19104, USA*
[c] *Kelley School of Business, Indiana University, Bloomington, IN 47405, USA*

Received 23 February 2005; accepted 8 May 2006
Available online 7 July 2006

Abstract

We propose a more complete conceptual framework for analysis of SME credit availability issues. In this framework, lending technologies are the key conduit through which government policies and national financial structures affect credit availability. We emphasize a causal chain from policy to financial structures, which affect the feasibility and profitability of different lending technologies. These technologies, in turn, have important effects on SME credit availability. Financial structures include the presence of different financial institution types and the conditions under which they operate. Lending technologies include several transactions technologies plus relationship lending. We argue that the framework implicit in most of the literature is oversimplified, neglects key elements of the chain, and often yields misleading conclusions. A common oversimplification is the treatment of transactions technologies as a homogeneous group, unsuitable for serving informationally opaque SMEs, and a frequent misleading conclusion is that large institutions are disadvantaged in lending to opaque SMEs.

☆ The opinions expressed do not necessarily reflect those of the Federal Reserve Board or its staff. The authors thank the anonymous referees, Zoltan Acs, Thorsten Beck, Lamont Black, Cesare Calari, Asli Demirguc-Kunt, Scott Frame, and Leora Klapper for valuable direction with the paper, and participants in the World Bank Conference on "Small and Medium Enterprises: Overcoming Growth Constraints" and the Australasian Finance and Banking Conference for helpful discussion.
* Corresponding author. Address: Board of Governors of the Federal Reserve System, Mail Stop 153, Federal Reserve Board, 20th and C Streets, Washington, DC 20551, USA. Tel.: +1 202 452 2903; fax: +1 202 452 5295.
E-mail addresses: aberger@frb.gov (A.N. Berger), gudell@indiana.edu (G.F. Udell).

0378-4266/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2006.05.008

JEL classification: G21; G28; G34; L11; F33

Keywords: SMEs; Banks; Relationship Lending; Governance, International

1. Introduction

The availability of external finance for small and medium enterprises (SMEs) is a topic of significant research interest to academics and an important issue to policy makers around the globe. The conceptual framework to which most of the current research literature adheres has been quite helpful in understanding the institutions and markets that provide funds to SMEs in developed and developing nations. This framework has also provided insights into the effects of policies that affect access to funding by creditworthy SMEs in these nations. However, we argue that the current framework is oversimplified, and neglects key elements of the financial system that affect SME credit availability.

We propose a more complete framework in which lending technologies play a key role as the conduit through which government policies and national financial structures affect SME credit availability. We define a lending technology as a unique combination of primary information source, screening and underwriting policies/procedures, loan contract structure, and monitoring strategies/mechanisms.

An important oversimplification in the current framework is the way that lending technologies are often categorized into two types: transactions lending that is based primarily on "hard" quantitative data and relationship lending, which is based significantly on "soft" qualitative information. Under this categorization, transactions lending is generally viewed as being focused on informationally transparent borrowers, while relationship lending is seen as used for opaque borrowers.

In our view, this characterization is fundamentally flawed. Transactions lending is not a single homogeneous lending technology. There are a number of distinct transactions technologies used by financial institutions, including financial statement lending, small business credit scoring, asset-based lending, factoring, fixed-asset lending, and leasing. While financial statement lending is focused on transparent borrowers, these other transactions technologies are all targeted to opaque borrowers. Recognition of this heterogeneity among transactions technologies and its impact on credit availability to opaque borrowers is often missing from the academic literature.

Our framework specifies a causal chain from government policies to a nation's financial institution structure and lending infrastructure. These financial structures, in turn, significantly affect the availability of funds to SMEs by determining the feasibility and profitability with which different lending technologies may be deployed. Financial institution structure refers to the market presence of and competition among different types of financial institutions and lending infrastructure refers to the rules and conditions that affect the ability of these institutions to lend. The extant research literature often neglects key elements of this causal chain, which may yield misleading research and policy conclusions.

To illustrate, consider the recent research on financial institution size, one dimension of financial institution structure. A common finding is that large institutions have a comparative advantage in transactions lending to SMEs based on hard information, while small institutions have a comparative advantage in relationship lending based on soft information. A policy implication that might at first blush seem reasonable is that the financial institution structure must include a substantial market share for small institutions to meet the demands of informationally opaque SMEs, since large institutions rely on hard information for their transactions lending technologies.

While the current conceptual framework is likely correct in associating large institutions with transactions technologies, the inference that large institutions are disadvantaged in lending to opaque SMEs is flawed. Large institutions deliver credit to many types of opaque SMEs through the transactions lending technologies that specifically address problems of informational opacity using hard information. For small business credit scoring, large institutions use hard information on the SME and/or its owner obtained from credit bureaus to infer future loan performance; for asset-based lending, these institutions use valuations of the assets pledged as collateral to evaluate repayment prospects; for factoring, they focus on the quality of the accounts receivable purchased; for fixed-asset lending and leasing, large institutions look to the valuations of the fixed assets that are pledged as collateral (fixed-asset lending) or directly owned by the institution (leasing). Thus, when informative financial statements are not available, institutions are often able to use other types of hard information to assess repayment prospects. Similar arguments apply to potentially misleading conclusions based on the current framework about other dimensions of a nation's financial structures and the policies that affect these structures.

Research on SME finance suffers from the problem that the lending technologies are usually not identified. This makes it difficult to test theories that relate financial structures to credit availability for different types of borrowers and to make policy assessments of which financial structures function best in supplying funds to creditworthy transparent and opaque SMEs. The limited findings from studies that identify lending technologies suggest that significant variation in the deployment of these technologies exists across nations – an institutional fact that is not explained by the current conceptual framework. For example, asset-based lending has a significant presence in only four nations, Australia, Canada, the UK, and the US. A goal for our framework is to try to explain a significant portion of the variation in the use of lending technologies with differences in national financial structures.

The effects of a nation's lending infrastructure on SME credit availability through determining the feasibility and profitability of deploying the different lending technologies is particularly under-researched in the literature. This infrastructure includes the information environment, the legal, judicial, and bankruptcy environments, the social environment, and the tax and regulatory environments in which financial institutions operate in a given nation. Lending infrastructures are quite heterogeneous across nations. We show how a nation's lending infrastructure affects the extent to which each of the individual lending technologies are employed in financing SMEs.

Section 2 briefly discusses each of the major lending technologies used for SMEs. Sections 3 and 4 focus on the financial institution structures and lending infrastructures of nations, respectively. We show how these structures may influence SME lending credit availability through affecting the feasibility and profitability with which the different lending technologies may be deployed. Section 5 provides some brief conclusions.

2. Lending technologies

We briefly define and describe each of the lending technologies, highlight its distinguishing features, and show how the technology addresses the opacity problem. Each technology is distinguished by a unique combination of the primary source of information, screening and underwriting policies/procedures, structure of the loan contracts, and monitoring strategies and mechanisms. In some cases, the technologies basically differ from one another in just one of these dimensions (e.g., fixed-asset lending and leasing primarily differ from each other on the contracting dimension – specifically, the ownership of the fixed assets). In other cases, lending technologies differ in multiple dimensions (e.g., small business credit scoring and asset-based lending differ in primary information sources and monitoring mechanisms).

In many cases, a secondary information source, screening/underwriting procedure, contract feature, or monitoring mechanism is used, but we distinguish the technologies based on the primary foundations of the lending decisions. Thus, a credit score may be used as secondary information or collateral may be used as a secondary source of repayment, but the lending technology would still be relationship lending if the lending decisions are primarily based on soft information gathered over the course of a relationship.[1]

2.1. Financial statement lending

Financial statement lending is a transactions technology based primarily on the strength of a borrower's financial statements. There are two requirements for this technology that depend on hard information. First, the borrower must have informative financial statements, such as audited statements prepared by a reputable accounting firm according to widely accepted accounting standards such as GAAP. Second, the borrower must have a strong financial condition as reflected in the financial ratios calculated from these statements. The loan contract that arises out of the analysis of these financial statements may reflect a variety of different contracting elements such as collateral and personal guarantees. However, under financial statement lending, the lender views the expected future cash flow of the SME as the primary source of repayment. Financial statement lending, unlike other lending technologies, is reserved for relatively informationally transparent firms.

2.2. Small business credit scoring

Small business credit scoring is a transactions technology based primarily on hard information about the SME's owner as well as the firm. The owner information is primarily personal consumer data obtained from consumer credit bureaus. This is combined with data on the SME collected by the financial institution and often from commercial credit bureaus. The data are entered into a loan performance prediction model, which yields a score, or summary statistic for the loan. The models are usually designed for credits up to $250,000, but many institutions use them only for credits up to $100,000. The technology is relatively new – it was not widely used in the US until the mid-1990s.

This technology may be applied to very opaque SMEs, given that much of the information is based on the personal history of the owner, rather than the SME. Consistent with this opacity, recent research finds that this technology is associated with credits under

[1] In some cases, the technologies may be most efficiently deployed in a particular organizational form or unit that is dedicated to that technology. The unit may be an entire financial institution (e.g., a leasing company offering only business leases), or a separate department, division, or subsidiary (e.g., an asset-based lending department of a commercial bank or finance company).

$100,000 that are rated as relatively risky, have high interest rates, and are often located outside of the banks' local markets (Frame et al., 2004; Berger et al., 2005b).[2]

2.3. Asset-based lending

Asset-based lending is a transactions lending technology in which financial institutions address the opacity problem by focusing on a subset of the firm's assets, which are pledged as collateral, as the primary source of repayment. This technology provides working capital financing secured primarily by accounts receivable and inventory.[3] The amount of credit extended is linked on a formula basis using hard data to a dynamically-managed estimation of the liquidation value of the assets used as collateral. The value of collateral is assessed daily in the case of accounts receivable, and typically weekly or monthly for inventory, and linked to the size of the credit available, so that the liquidation value of the collateral always exceeds the credit exposure (Udell, 2004).[4]

The use of collateral itself, however, does not distinguish asset-based lending from the other lending technologies. The pledging of accounts receivable and inventory is often associated with financial statement lending, relationship lending, and credit scoring, where collateral is used a secondary source of repayment. Under asset-based lending, in contrast, the extension of credit is primarily based on the value of the collateral, rather than the overall creditworthiness of the firm.

2.4. Factoring

Factoring involves the purchase of accounts receivable by a "lender" known as a factor. As in asset-based lending, factoring focuses on the value of an underlying asset, rather than the overall value/risk of the firm. Factoring is similar to asset-based lending, but there are three important distinctions. First, factoring only involves the financing of accounts receivable, unlike asset-based lending which also involves financing inventory. Second, under factoring, the underlying asset is sold to the "lender." Third, factoring is essentially a bundle of three financial services: a financing component, a credit component and a collections component. Under most factoring arrangements, the borrower outsources its credit and collections activities in addition to obtaining financing. Factoring is a transactions technology because the underwriting process is based on hard information about the value of a "borrower's" accounts receivable. Factoring addresses the

[2] See Berger and Frame (2006) for more details on small business credit scoring and its effects on credit availability.

[3] There is relatively little empirical evidence on asset-based finance. One study finds evidence consistent with practitioner and conventional wisdom that asset-based finance is associated with riskier borrowers (Klapper, 1998). Another study of very large credits finds evidence that commercial finance company credits (which are exclusively asset-based) are riskier than commercial bank credits (which can be either asset-based or not), but did not find evidence associating opacity with asset-based lending using standard opacity measures (Carey et al., 1998).

[4] For the least creditworthy asset-based borrowers, lenders often establish special checking accounts into which all collections of receivables must be deposited, reducing potential diversion of these funds by the borrower (Mester et al., forthcoming). In addition, monitoring cash flows on deposit accounts may provide valuable hard information that may be used in conjunction with any of the lending technologies, provided that all the firms' checking accounts are consolidated at the lending institution (Nakamura, 1993).

opacity problem by focusing primarily on the quality of the obligor, rather than the "borrower."[5]

The use of factoring varies widely across countries. For example, the ratio of the volume of factoring to GDP in 2002 was 11.9% in Italy, but only 0.9% in Switzerland (Bakker et al., 2004).

2.5. Fixed-asset lending

Fixed-asset lending technologies involve lending against assets that are long-lived and are not sold in the normal course of business (e.g., equipment, motor vehicles, or real estate).[6] Like asset-based lending – which is based on accounts receivable and inventory – the underlying assets in fixed-asset lending are pledged to the lender as collateral. However, unlike accounts receivable and inventory, the pledged assets are virtually always uniquely identified by a serial number or a deed. The long life and unique identification of fixed assets leads to very different underwriting processes, contract structures, and monitoring mechanisms. At the underwriting stage, the focus is on assessing the market value of the asset – for equipment and real estate, this is often in the form of a formal appraisal. The contract structure typically specifies an initial loan-to-value ratio less than one. It also typically involves setting a loan amortization schedule with a final maturity less than the lifespan of the asset. The schedule insures that the ratio of the outstanding loan balance to the liquidation value of the asset is less than one over the life of the loan. This contract structure also feeds back to the underwriting process, where the primary financial analysis focuses on coverage ratios that measure a firm's ability to meet the amortization schedule (i.e., debt service).[7] Unlike monitoring asset-based loans, the monitoring of collateral (i.e., ownership of collateral by the borrower) is not problematic in fixed-asset lending. The borrower can only sell a fixed asset by transfer of title, which can only occur if the lender agrees to release the lien. Monitoring the borrower's ability to pay (i.e., its cash flow) is tied to the observation of timely repayment as specified by the amortization schedule. Failure to meet a required payment signals inadequate cash flow and triggers a default on the loan.

2.6. Leasing

Leasing involves the purchase of fixed assets by a "lender" known as a lessor. Leasing is a very common method of financing equipment, motor vehicles, and real estate in many countries by both banks and other institutions. The lessor purchases the fixed assets and then simultaneously enters into a rental contract with the lessee (the "borrower") that

[5] "Reverse factoring" is a recent innovation that has been applied in developing economies. Under reverse factoring, a factor enters into an agreement with a large company (e.g., Wal-Mart) that is purchasing goods from a large number of small suppliers. The factor agrees to finance any of the receivables of this large company generated by invoices from these small suppliers. Reverse factoring differs from conventional factoring in that the primary business arrangement is between the seller of the goods and the factor, as opposed to the supplier of the goods and the factor (Klapper, 2006).

[6] In SME lending, fixed-asset lending can be associated with personal assets provided by the firm's owner as well as the firm's assets. A common example is an entrepreneur pledging a personal residence as collateral for a business loan.

[7] Coverage ratios are calculated by dividing a measure of cash flow by a measure of debt service.

specifies the payment schedule. The contract often contains an option whereby the lessee can purchase the assets at the end of the lease at a pre-specified price.

Leasing is a transactions technology because underwriting is substantially based on hard information about the value of the underlying asset, analogous to asset-based lending, factoring, and fixed-asset lending. Like these other technologies, leasing can be used to provide financing to opaque firms because the underwriting decision is primarily based on the value of the asset being leased. It has also been shown that leasing can mitigate an adverse selection problem, either in the used equipment market by encouraging a higher quality of product sold "off-lease" (i.e., sold by lessors when the purchase option is not exercised) or in the new product market (Chemmanur and Yan, 2000; Hendel and Lizzeri, 2002; Gilligan, 2004).[8]

2.7. Relationship lending

Under relationship lending, the financial institution relies primarily on soft information gathered through contact over time with the SME, its owner and the local community to address the opacity problem. This information is acquired in large part by the loan officer through direct contact with the borrower and through observing the SME's performance on all dimensions of its banking relationship. This soft information may also include an assessment of the future prospects of the SME garnered from past communications with SME's suppliers, customers, or neighboring businesses (Petersen and Rajan, 1994; Berger and Udell, 1995; Degryse and Cayseele, 2000). As noted, this soft information may often remain proprietary to the loan officer because it is not easily observed, verified, or transmitted to others.

2.8. Trade credit

The importance of trade credit to SME financing is a compelling reason to include it as a lending technology, although it is not delivered by financial institutions. Trade credit represented one-third of the debt of US SMEs as of 1998, nearly that extended by commercial banks (Robb, 2002). Trade credit may be particularly important in economies with weak financial systems, where industries with higher dependence on trade credit exhibit higher growth rates (Demirguc-Kunt and Maksimovic, 2002; Fisman and Love, 2003).

It is not straightforward to classify trade credit as either a transactions based or a relationship based technology. Many of the procedures and processes associated with the other lending technologies are used in underwriting trade credit. For example, credit scoring and similar quantitative techniques have long been a part of the underwriting process used by credit managers. For larger accounts, financial statements are analyzed as part of the underwriting process. However, soft information and mutual trust play a role in some trade credit underwriting, similar to relationship lending.

Researchers have suggested comparative advantages for product sellers in funding, production/inventory management, price discrimination, or product quality guarantees. Some have also suggested that suppliers may have an informational advantage over other

[8] Leasing may also provide for an optimal sharing of tax benefits when there is a differential tax rate between the lessor and the lessee (e.g., Graham et al., 1998).

lenders in evaluating their customers' ability to pay, solving incentive problems effectively, repossessing and reselling goods in the event of default, or withholding future supplies (see Petersen and Rajan, 1997 for a summary of this research). More recent evidence suggests additional advantages of trade credit. These include a signaling role in developing economies, reduced incentive to strategically default, an informational advantage in lending to less creditworthy firms when trade creditors are given a security interest, and an advantage in financing non-standardized products (Cook, 1999; Cunat, 2003; Frank and Maksimovic, 2003; Burkart and Ellingsen, 2004; Burkart et al., 2005; Chan et al., 2006).

3. Financial institution structure

In this section, we focus on the effects of a nation's financial institution structure on the feasibility and profitability with which the different lending technologies can be deployed to fund SMEs. The research literature provides a considerable amount of evidence on the effects of financial institution structure on SME lending, but the findings rarely go beyond the distinction between transactions lending technologies versus relationship lending to parse among the different transactions technologies. We briefly review the findings with regard to the comparative advantages of large versus small institutions, foreign versus domestically-owned institutions, state versus privately-owned institutions, and market competition.

3.1. Large versus small institutions

For a number of reasons, large institutions may have a comparative advantage in transactions lending and small institutions may have an advantage in relationship lending. Large institutions may be able to take advantage of economies of scale in the processing of hard information, but be relatively poor at processing soft information because it is difficult to quantify and transmit through the communication channels of large organizations (e.g., Stein, 2002). Under relationship lending, there may be agency problems created within the financial institution because the loan officer that has direct contact over time with the SME is the repository of soft information that cannot be easily communicated within the financial institution. This may give a comparative advantage in relationship lending to small institutions with fewer layers of management (e.g., Berger and Udell, 2002). Finally, large institutions may be disadvantaged at relationship lending because of Williamson-type (Williamson, 1988) organizational diseconomies associated with also providing transactions loans and other wholesale services.

The empirical literature generally does not identify the lending technologies, but in some cases draws conclusions from the characteristics of the SME borrowers and their relationships with financial institutions. In most cases, the research is based on data from US banks and SMEs. Large institutions are found to lend to larger, older SMEs with stronger financial ratios, and small institutions are found to rely more on soft information and lend to SMEs with which they have stronger relationships (e.g., Haynes et al., 1999; Cole et al., 2004; Scott, 2004; Berger et al., 2005c). The strong financial ratios and SME size and age findings are consistent with a comparative advantage for large institutions in using financial statement lending based on financial ratios for relatively transparent SMEs (which tend to be larger and older), given that financial statement lending is the only lending technology generally used for transparent borrowers. The strong relationships and use of soft information by small institutions is consistent with a comparative advantage for small institutions in relationship lending. There is very little evidence on comparative advantages by financial institution size in the other transactions lending technologies, although it is likely that large institutions have comparative advantages of differing degrees in the other transactions technologies due to economies of scale in processing hard information.

A limited amount of empirical evidence also addresses the issue of a general comparative advantage of large versus small institutions in SME lending and the policy question of whether a sizeable presence of small institutions is necessary to promote credit availability for SMEs. One study finds that the likelihood that an SME has a line of credit from a large bank versus a small bank is roughly proportional to the deposit market presence of large versus small banks (Berger et al., 2006). This is not consistent with a general comparative advantage by financial institutions size and suggests that a strong presence of small institutions may not be needed for general SME credit availability. This study also found few significant differences in the opacity or other characteristics of small businesses that borrow from large versus small banks. These two findings are consistent with the hypothesis that large institutions using various transactions technologies are just as able to supply SME credit to opaque SMEs as small institutions using relationship lending.

Another study found that the local market shares of large and small US banks have relatively little association with SME credit availability, but did not account for the size of the institutions making the loans (Jayaratne and Wolken, 1999). This study is also consistent with the hypothesis that large institutions are not disadvantaged in SME lending and that a strong presence of small institutions may not be needed for general SME credit availability. An alternative hypothesis to explain this finding is that large institutions may be generally disadvantaged at SME lending, but that a sufficient market presence of small institutions efficiently arises endogenously when needed to provide adequate SME credit. Supporting this second hypothesis, several studies find that large institutions reduce their SME lending after mergers and acquisitions, but that other banks in the same local markets appear to respond by increasing their supply of SME credit substantially (e.g., Berger et al., 1998, 2001a; Avery and Samolyk, 2004). As well, new small banks are often created in these markets that supply additional SME credit (Berger et al., 2004a).

Thus, the research suggests that concern about the policy issue of whether a sizeable presence of small institutions is needed for SME credit availability in US markets may not be warranted. However, these findings may not apply to other nations because of differences in financial institution structure and lending infrastructure that may limit the use of some of the lending technologies. In an international comparison, greater market shares for small banks are associated with higher SME employment and more overall bank lending in both developed and developing nations (Berger et al., 2004b).

3.2. Foreign-owned versus domestically-owned institutions

Foreign-owned institutions may have a comparative advantage in transactions lending and a disadvantage in relationship lending in part because these institutions are typically large. In developing nations, foreign institutions headquartered in developed nations may have an additional advantage in transactions lending because of access to better information technologies for collecting and assessing hard information.[9] Foreign institutions may also face additional hurdles in relationship lending because of difficulties in processing and transmitting soft information over greater distances, through more managerial layers, and coping with multiple economic, cultural, language, and regulatory environments (e.g., Buch, 2003).

There is little empirical evidence on SME lending by foreign-owned institutions in developed nations. Some research finds that these institutions tend to have a wholesale orientation (e.g., DeYoung and Nolle, 1996), and in some cases tend to specialize in serving multinational corporations headquartered in their home nation (e.g., Goldberg and Saunders, 1981). Presumably, these institutions use transactions technologies applied to hard information.

There is more evidence on foreign-owned institutions and SME credit availability in developing nations. In most of the studies, foreign banks are associated with greater SME credit availability (e.g., Beck et al., 2004a; Berger et al., 2004b; Clarke et al., 2005a). However, some find that foreign-owned banks may have more difficulty in supplying credit to SMEs than to large firms (e.g., Berger et al., 2001b; Mian, forthcoming). This relative preference for large firms over SMEs has also been found for foreign lenders entering formerly socialist countries (Giannetti and Ongena, 2005). Again, the lending technologies are generally not identified. Although foreign institutions almost surely use transactions technologies, it is usually not known which of these technologies are employed or the opacity of the borrowers served.[10]

3.3. State-owned versus privately-owned institutions

State-owned institutions may be expected to have a comparative advantage in transactions lending and a disadvantage in relationship lending because these institutions are typically large. Other arguments regarding the ability of these institutions to supply funds to creditworthy SMEs through any lending technology may also apply. State-owned institutions generally operate with government subsidies and often have mandates to supply additional credit to SMEs in general, or to those in specific industries, sectors, or regions. Although this might improve funding to creditworthy SMEs in principle, it could have the opposite effect in practice because these institutions may be inefficient due to a lack of market discipline. As well, much of their SME funding may be to SMEs that are not creditworthy because the lending mandates do not necessarily require that the funding be used to finance positive net present value projects, or that loans be repaid at market rates. Some of the funds may also be subsidized or directed for political purposes, rather than economic ends (e.g., Cole, 2004; Sapienza, 2004). State-owned institutions may also deploy

[9] For example, some foreign-owned institutions use a simplified form of small business credit scoring to lend to SMEs in developing nations based on the SMEs industry or provide home-nation training for loan officers stationed in developing nations (Berger et al., 2004b).

[10] There is also research on the efficiency of foreign versus domestic institutions, but the efficiency differences may or not be linked to lending technologies and SME credits. The research in developed nations suggests that foreign institutions are less efficient on average (e.g., DeYoung and Nolle, 1996; Berger et al., 2000). However, in developing nations, foreign-owned banks usually appear to be more profitable and efficient than domestically-owned banks on average (e.g., Claessens et al., 2001; Martinez Peria and Mody, 2004), although one study finds roughly equal performance after controlling for a number of different types of governance and governance change (Berger et al., 2005a).

relatively weak monitoring strategies and/or refrain from aggressive collection procedures as part of their mandate to subsidize targeted borrowers or because of the lack of market discipline. In nations with substantial state-owned banking sectors, there may also be significant spillover effects that discourage privately-owned or foreign-owned institutions from SME lending due to a "crowding out" effect from subsidized loans from state-owned institutions.

The empirical findings are generally consistent with unfavorable effects of state ownership. Studies of general performance typically find that individual state-owned banks are relatively inefficient and that large shares of state bank ownership are typically associated with unfavorable macroeconomic consequences and less developed financial and economic systems (e.g., La Porta et al., 2002; Barth et al., 2004; Berger et al., 2004b). Some evidence also suggests that less SME credit is available in nations with large market shares for state-owned banks (e.g., Beck et al., 2004a; Berger et al., 2004b). As well, nonperforming loan ratios at state-owned banks tend to be very high, consistent with lending based on negative net present value projects, weak monitoring, and/or lack of aggressive collection procedures (e.g., Berger et al., 2005a). Studies of the effects of bank privatization in both developed nations and developing nations typically find improvements in performance following the elimination of state ownership (e.g., Clarke et al., 2005b). Similar to the case for foreign-owned institutions, state-owned institutions likely use transactions technologies, but the technologies are generally not identified in the research.[11]

3.4. Market competition

Under the traditional structure-conduct-performance (SCP) hypothesis, market power reduces credit access through any lending technology. Institutions with more market power may charge high rates or fees on loans; have tight credit standards; and/or be less aggressive in finding or serving creditworthy SMEs so managers can take advantage of a "quiet life." An alternative hypothesis suggests that for one of the lending technologies, relationship lending, market power may be associated with greater access to credit for SMEs. Market power may encourage institutions to invest in lending relationships because the SMEs are less likely to find alternative sources of credit in the future. Market power helps the institution enforce a long-term implicit contract in which the borrower receives a subsidized interest rate in the short term, and then compensates the institution by paying a higher-than-competitive rate in a later period (Sharpe, 1990; Petersen and Rajan, 1995). However, under a different theoretical model of relationship lending, greater concentration may be associated with less credit availability using this technology (Boot and Thakor, 2000).

Thus, economic theory offers conflicting empirical predictions that arise out of different theoretical models about one of the lending technologies, relationship lending. Without separately identifying relationship lending from the other technologies, the empirical predictions for access to credit for SMEs are not clear For example, market power could increase credit availability for some SMEs through a positive effect on relationship lending, but may decrease availability for other SMEs that are more suited to one of the

[11] There are exceptions to the measured unfavorable effects of state ownership – cases in which state-owned institutions have eliminated government subsidies and appear to operate relatively efficiently and provide significant SME credit (e.g., Townsend and Yaron, 2001).

transactions lending technologies. Thus, the effect of market power on overall SME credit availability may go either way, depending on the strength of the different hypotheses and the extent to which the different lending technologies are employed.

A number of studies have looked at measures of SME credit availability, activity, and general economic performance and their association with indicators of market power such as concentration and regulatory restrictions on entry and competition. The empirical results are mixed, with some studies finding generally unfavorable effects from market power (e.g., Elsas, 2005; Karceski et al., 2005; Cetorelli and Strahan, 2006), and others finding favorable effects (e.g., Petersen and Rajan, 1995; Cetorelli and Gambera, 2001; Bonaccorsi di Patti and Dell'Ariccia, 2004; Cetorelli, 2004).[12] One study found that the results depend on how market power is measured with concentration measures and the Lerner index generally showing conflicting results (Carbó-Valverde et al., 2006).

4. The lending infrastructure

In this section, we turn our attention to the lending infrastructures of nations and how they affect the feasibility and profitability of using the different lending technologies in SME financing. The lending infrastructure includes the information environment, the legal, judicial and bankruptcy environments, the social environment, and the tax and regulatory environments. All of these elements may affect SME credit availability by influencing the extent to which the different lending technologies may be legally and profitably employed. The final element, the regulatory environment, may also restrict SME credit availability by constraining the financial institution structure.

4.1. The information environment

An important aspect of the information environment is the accounting infrastructure. Strong accounting standards and credible independent accounting firms are necessary conditions for informative financial statements, which are key to the financial statement lending technology. These are also important for many components of loan contracting associated with financial statement lending and some of the other lending technologies to a lesser extent. For example, covenants based on financial ratios are not feasible if the ratios calculated from the financial statements are not reliable. Indices of global accounting standards indicate considerable variation across countries both between developed and developing economies and among developed economies (e.g., La Porta et al., 1998).

Another important aspect of the information environment is the sharing of information. Commercial and consumer credit bureaus provide formal organizational mechanisms for the exchange of payment performance data. Credit bureaus have also been found to reduce the cost and time to process loans and the level of defaults (Miller, 2003). Commercial credit bureau data have also been shown to have power in predicting firm failure

[12] Some studies also found that higher measured market power is associated with higher SME loan interest rates (e.g., Hannan, 1991; Berger et al., 2006). Although this finding may appear to support the SCP hypothesis, it may also be consistent with the alternative hypothesis of an expansion of relationship lending if relationship loans tend to have higher interest rates on average than transactions loans (which may or may not be the case). Again, much of the difficulty in interpreting the results arises because the lending technologies are not identified.

beyond financial ratios and other descriptive information about the firm (Kallberg and Udell, 2003).

There is considerable variation across countries in terms of the existence of credit bureaus, whether they are publicly- or privately-owned, and the coverage of available information (Miller, 2003). Empirical evidence suggests a statistically important link between the existence of third-party information exchanges and credit availability (Pagano and Jappelli, 1993; Padilla and Pagano, 1999; Love and Mylenko, 2003). Specifically, countries with stronger formal information sharing exhibit greater bank lending relative to GNP and country-level credit risk is negatively correlated with measures of formal information sharing (Jappelli and Pagano, 2002).

Credit bureau information, where available, is used in conjunction with all of the lending technologies, but it is a driving component of one of the lending technologies, small business credit scoring. It is necessary to have a large database on SME loan performance and the variables used to predict that performance in order to estimate a credible credit scoring model. In the US, most banks use external vendors to create the scores, and these vendors rely on a strong information environment in which credit bureaus share consumer and business information and financial institutions share their loan performance data to estimate the model. Some large institutions are also able to generate proprietary models based on their own loan performance experience, but still need information from credit bureaus for the exogenous variables in their models. The strong information environment requirements may help explain why small business credit scoring appears to be used on a widespread basis only in a limited number of developed nations (Miller and Rojas, 2004).[13]

The evidence also indicates that access to historical credit information such as business credit bureaus is positively related to the presence of factoring (Klapper, 2006). Under certain circumstances factoring can also work well even in weak domestic information environments if the receivables are from obligors located in strong information environments. For example, the receivables of an Estonian firm whose customers are located in Germany may qualify for factoring because the factor can efficiently assess the creditworthiness of the German account obligors (Bakker et al., 2004).

4.2. The legal, judicial, and bankruptcy environments

A country's legal, judicial, and bankruptcy environments significantly influence the context in which loan contracting is conducted. The legal environment that affects business lending consists of the commercial laws that specify the property rights associated with a commercial transaction. The judicial and bankruptcy environments determine how well these laws are enforced in commercial disputes and in bankruptcy resolutions. This enforceability, in turn, determines the confidence of contracting parties in financial contracts. Collectively, these features constitute the rule of law as it relates to the extension of credit. Countries differ significantly on this dimension: for some, commercial laws are unambiguous and conducive to commercial transactions and enforcement is predictable; for others commercial law is ambiguous and incomplete, enforcement is problematic,

[13] There are some efforts to expand the use of this lending technology in developing nations. The World Bank and Fair, Isaac have recently formed a joint venture to determine the feasibility of developing a pooled data small business credit scoring model using data from Brazil, Columbia, and Mexico (Miller and Rojas, 2004).

and criminal and racketeering behavior block the creation of new businesses, undermine existing ones, and deter foreign investment (EBRD, 2003).

Empirical studies have shown that firms in countries with greater financial development and stronger property rights have increased levels of investment funded by external finance. Firms in countries with weaker financial development and property rights, in contrast, are more likely to rely on potentially less efficient financing from development banks, the government, or informal sources (Beck et al., 2004b). Smaller firms may be particularly affected. One study found that the effect of financial, legal, and corruption problems consistently constrained the growth of smaller firms more than larger firms in a cross-country analysis (Beck et al., 2005).

Commercial laws and their enforcement of these laws also affect the ability of banks to deploy specific contracting elements that can be used to address informational opacity problems. Specifically, they can affect the deployment of contracting elements such as covenants, maturity, collateral, and personal commitments that have been shown to mitigate adverse selection and moral hazard problems (e.g., Chan and Kanatas, 1985; Berlin and Loeys, 1988; Sharpe, 1990; Berkowitz and White, 2004).

The commercial law on security interests (collateral liens) in a nation, for example, is important in determining the efficacy of collateral in a loan contract, the essential component of the asset-based lending and fixed-asset lending technologies. Key issues include whether a country's commercial law clearly defines how a collateral lien can be perfected, how collateral priority is determined, and how notification of a lien is made. At one extreme are countries such as the US that have a well-developed set of commercial laws that include blanket filings on accounts receivable and inventory (Article 9 of the Uniform Commercial Code) and well-defined electronic registration systems that temporally define lien filings. At the other extreme are countries with, at best, recently implemented commercial law that leaves collateral priority and notification ambiguous. Empirical evidence suggests that the legal efficacy of collateral – in particular, the ability of lenders to seize collateral in the event of default – is associated with increased use of collateral (Qian and Strahan, 2005). Although progress is being made on this dimension, many of these countries – such as those in Eastern Europe – are still deficient in key areas such as the scope of assets that can be secured, registration and filing, priority, and enforcement (EBRD, 2003).[14]

The efficiency of the judicial and bankruptcy systems are also critical to credit availability. Recent theoretical and empirical research suggests that judicial inefficiency (i.e., high-cost judicial procedures) is associated with decreased access to credit (Jappelli et al., 2005). It has also been shown that countries with greater legal procedural formalism – typically associated with civil law countries – take longer to enforce some types of financial contracts (Djankov et al., 2003). The Czech Republic offers an illustration of this problem. The cost of enforcing commercial claims was considerably higher in the Czech Republic than in five other transition economies that have joined the EU, the 15 non-accession EU countries, and the US. Pursuing claims in a bankruptcy environment is also quite difficult in this country (World Bank, 2003).

The length of time in bankruptcy is a particularly important dimension of efficiency. Also important is the degree to which the application of bankruptcy procedures are associated

[14] In many of these countries, lenders cannot file a single lien on all currently existing and future accounts receivable and inventory and instead must identify by invoice number and serial number (where applicable) each receivable and inventory item as it is generated. This effectively renders asset-based lending infeasible.

with adherence to absolute priority. The power of collateral *ex ante* ultimately depends on the whether the priority rights of secured lenders are upheld in bankruptcy *ex post*. Details of the laws that are often missed in academic analyses can be extremely important. For example, the rights of secured lenders in the US may at first seem relatively weak because an automatic stay is immediately invoked upon acceptance of a bankruptcy petition by the bankruptcy court.[15] The automatic stay prevents all creditors from collecting payments from the bankrupt firm and otherwise enforcing their financial claims. However, under US bankruptcy law, the judge is required to preserve the collateral claim of secured creditors and to give them "adequate protection" if the collateral or its proceeds are denied to the secured lender. That is, the bankruptcy judge is obligated to preserve the value of a secured lender's claim. Moreover, petitions by secured lenders for a waiver from the automatic stay are often filed at the time of the bankruptcy petition and approved by the judge in the case of specific collateral classes, such as accounts receivable (Udell, 2004).

The fact that asset-based lending has a significant presence in only four nations suggests that these legal, judicial, and bankruptcy environmental conditions represent significant hurdles. Nevertheless, in the countries where asset-based lending exists, it appears to be quite important. In the US, for example, the stock of total asset-based loans is about \$300 billion (Commercial Finance Association, 2003), compared to the stock of commercial and industrial loans in the US of about \$900 billion (inclusive of bank asset-based loans).

By the same token that asset-based lending and fixed-asset lending depend on a strong lending infrastructure, a weak lending infrastructure may foster the use of factoring and leasing. This is because under factoring or leasing, the underlying assets – accounts receivable or fixed assets, respectively – are removed from the "borrower" (and the borrower's bankruptcy estate) and owned by the "lender," rather than being pledged as collateral. Thus, a lending infrastructure that does not support the use of collateral may encourage the use of these alternatives where the lender owns the asset. Consistent with this, there is evidence that factoring is more important in countries with weak contract enforcement (Klapper, 2006). Analyses of transitioning former socialist nations also suggest that weak enforcement may encourage leasing (World Bank, 2000, 2002).

Similarly, a weak lending infrastructure may encourage the substitution of relationship lending for some of the transactions lending technologies. The use of implicit contracts in relationship lending may substitute for the explicit contracts of transactions lending when explicit contracts cannot be easily enforced. While this issue has not been examined empirically, there is evidence that legal environments may be associated with relationship-based finance in the context of bank-based versus capital-markets-based financial systems. That is, weaker legal systems may have more bank-based rather than capital-markets-based financing (Rajan and Zingales, 2003; Ergungor, 2004). By extension, this argument may apply to relationship lending versus the transactions technologies.

4.3. The social environment

The social environment may also affect SME credit availability. Evidence suggests that the level of social capital and trust may be important in facilitating the writing and

[15] This was one of the reasons why the US received a low rating in academic analyses of business contracting environments (e.g., La Porta et al., 1998).

enforcement of financial contracts. Social capital, as proxied by electoral participation, has been found to be significant in explaining regional differences in some measures of entrepreneurial activity, such as number of firms and new firm entry (Guiso et al., 2004). Common language may also help develop mutual trust and facilitate relationship building. It has also been found that cultural differences across countries are associated with differences in the level of investor protection (Stulz and Williamson, 2003).

The greatest impact of the social environment is likely on relationship lending because social norms, religion, and culture may have the most effect on the production of soft information and the ability of banks to use this information to forge relationships. The social environment could also affect the production and use of hard information used in the transactions technologies, but likely to a lesser degree.

4.4. The tax and regulatory environments

The tax and regulatory environments may have significant effects on SME credit availability. For example, stamp taxes on factored invoices and certain types of value-added taxes can have negative impacts on factoring. Changes in capital regulations and tougher bank supervision in the US are also often cited as contributing to the reduction in supply of business credit or "credit crunch" in the US in the early 1990s (e.g., Berger and Udell, 1994; Peek and Rosengren, 1995). The implementation of the Basel II capital requirements may also have an impact in the future by imposing a different implicit tax on SME lending and favoring some lending technologies over others for some financial institutions. For example, large banking organizations may have more favorable capital requirements if they manage SME loans as part of a pool, which may favor some of the transactions technologies over relationship lending (Altman and Sabato, 2005; Berger, 2006).[16]

The regulatory environment may also affect SME credit availability by constraining the financial institution structure. Government policies often affect the entry of different types of financial institutions, their market shares, their abilities to compete, and their corporate governance structure. In many parts of the world, the removal of geographic and product restrictions has resulted in significant consolidation within the banking industry and between banks and other types of financial institutions. In the EU, the single banking license and other parts of the Single Market Programme appeared to spur considerable financial institution consolidation within nations and somewhat less activity across international borders within the EU. In the US, removal of state geographic restrictions and the Riegle-Neal Act clearly led to considerable consolidation within the US banking industry, although the Gramm-Leach-Bliley Act did not result in much additional consolidation between banking organizations and other types of financial institutions. However, as discussed in Section 2, the effects of bank size structure on SME credit availability are ambiguous.

Government policies that restrict foreign entry may have large effects on SME credit availability, given the findings that larger market shares for foreign-owned banks are often

[16] In the US, the largest banks are generally expected to have lower capital requirements on many SME loans, while in other nations, banks of all sizes may have lower requirements on many SME loans. To receive the favorable retail loan categorization, the exposure to the SME must be under €1 million, and the credit must be managed as a retail exposure as part of a pool or be guaranteed by an individual. For SME loans not classified as retail, there may also be a favorable treatment in some circumstances for loans to firms with under €50 million in sales.

associated with greater SME credit availability in developing nations. Other research has also found that restrictions on foreign-bank entry may be more strongly linked to bank performance than the market presence of foreign-owned banks (Levine, 2003), which may suggest that these restrictions have particularly strong effects in limiting competition, with potential consequences for SME customers. As well, restrictions on foreign institutions may limit the efficient use of some of the transactions lending technologies in developing nations, given the likely advantage of these institutions in collecting and processing hard information.

Finally, government policies with respect to state ownership of financial institutions clearly have important effects on credit availability. State ownership is generally found to have significant negative effects on SME credit availability, with some reversal of these effects after privatization. State-owned institutions likely use transactions lending technologies, but the evidence does not suggest which ones these institutions use or that they have advantages in any of the technologies.

5. Conclusions

We offer a more complete conceptual framework for thinking about the research and policy issues surrounding the availability of credit to SMEs in various circumstances around the globe. We emphasize a causal chain in which the lending technologies provide the crucial link between government policies and financial structures on the one hand, and SME credit availability on the other hand. At the top of the chain, government policies affect a nation's financial institution structure and lending infrastructure. That is, policies help influence the market shares and competitive conditions for large versus small, foreign versus domestic, and state-owned versus private financial institutions (financial institution structure), and the information, legal, judicial, bankruptcy, social, tax, and regulatory environments in which these institutions operate (lending infrastructure). These financial structures then help determine the feasibility and profitability with which the different lending technologies can be deployed to fund SMEs. The financial institution structure affects the use of technologies because the institution types have comparative advantages in different lending technologies. The lending infrastructure affects the legality and profitability of the lending technologies.

At the bottom of the causal chain, the lending technologies have important effects on the access to credit for creditworthy transparent and opaque SMEs. The different technologies – financial statement lending, small business credit scoring, asset-based lending, factoring, fixed-asset lending, leasing, relationship lending, and trade credit – each involves a different combination of primary information source, screening and underwriting policies/procedures, loan contract structure, and monitoring strategies/mechanisms. The choice of lending technology for a specific creditworthy SME depends on the sources of information available for that firm, as well as the adaptability and appropriateness of the various screening, underwriting, contracting, and monitoring techniques dealing with the firm in its environment.

We argue that the framework in the extant research literature is oversimplified and often neglects key elements of this causal chain. In some cases, the literature on SME credit availability has not accounted for or controlled for the presence of alternative lending technologies. A common oversimplification is to treat all transactions lending technologies as a homogeneous group that is not suitable for lending to informationally opaque

SMEs. In contrast, we argue that most of the transactions technologies are designed to use hard information other than financial statements to underwrite loans to opaque SMEs. These oversimplifications have often led to misleading conclusions regarding how policies may affect SME credit availability through their impacts on financial institution structure and lending infrastructure.

We also review much of the extant research on SME credit availability through the lens of this more complete framework, yielding several conclusions. First, the findings argue against drawing simplistic conclusions from the extant research. For example, the finding that large financial institutions have a comparative advantage in transactions lending technologies and comparative disadvantage in relationship lending does not necessarily imply that large institutions are disadvantaged in providing credit to informationally opaque SMEs. On the contrary, some of the transactions lending technologies used by large institutions are well-suited for funding opaque SMEs. Similarly, there is no simple answer to the policy question of whether a sizeable presence of small institutions is needed for overall SME credit availability. A limited amount of research using US data suggests little concern on this issue, but other research suggests that small institution presence may be more important in other nations because differences in financial structures that may limit the use of some lending technologies.

Second, the results make a strong case for taking into account the presence of foreign- and state-owned institutions, as well the presence of large and small institutions and conventional measures of financial institution competition, particularly when analyzing developing nations. All of these elements of financial institution structure may affect SME credit availability through comparative advantages in the different lending technologies. In particular, a greater presence of foreign-owned institutions and a lesser presence of state-owned institutions is likely to be associated with significantly higher SME credit availability in developing nations because foreign-owned institutions appear to have advantages in some of the lending technologies, and state-owned institutions appear to be generally disadvantaged.

Third, our investigation strongly suggests that lending infrastructures have important effects on SME credit availability. "Better" lending infrastructures may significantly improve SME credit availability through facilitating the use of the various lending technologies. As examples, better accounting standards may help spur the use of financial statement lending, better commercial laws on security interests may facilitate the use of asset-based lending and fixed-asset lending, and greater sharing of information may help facilitate the use of small business credit scoring, although other parts of the lending infrastructure must also be in place for these technologies to be legally and profitably employed. Similarly, "worse" lending infrastructures may appreciably reduce SME credit availability. This may occur if a restrictive regulatory environment constrains the financial institution structure, preventing some types of financial institutions from capitalizing on their comparative advantages in specific lending technologies. The evidence suggests that these effects may be quite strong, particularly when governments restrict the entry of foreign financial institutions and/or maintain large market shares for state-owned institutions.

For a framework to be useful, it must have clear testable implications, so that the proposed paradigm may be supported or refuted by the data. Our framework postulates how financial structures affect the feasibility and profitability of the different lending technologies, and the effects of these technologies on SME credit availability. This framework has a number of clear testable implications for the links among a nation's financial institution structure, financial infrastructure, lending technologies employed, and SME credit availability.

To the extent that (a) differences in financial structures across nations explain a significant portion of the observed variation in the use of different lending technologies; and (b) these different technologies explain a significant portion of the observed variance in SME credit availability, the framework would be supported. If differences in the use of the lending technologies are not strongly related to financial institution structure or lending infrastructure, or if the use of the technologies is not strongly associated with SME credit availability, the framework would be refuted.

While comprehensive tests are best, individual elements of the framework may also be tested. The framework predicts specific associations among the individual elements of the financial institution structure, lending infrastructure, technologies, and SME credit availability. For example, the paradigm predicts that the presence of large banks and quality credit bureaus may facilitate the use of small business credit scoring, which, in turn, may be efficiently used to lend to some opaque SMEs. Findings confirming these individual effects would support the framework, while findings of no effects of bank size or credit bureaus on small business credit scoring, or no effect of small business credit scoring on lending to opaque SMEs would tend to refute the framework.

A key issue in testing the framework is the identification of the lending technologies. One solution is for researchers to use the existing data sets to separate out and identify specific lending technologies that often have not been identified in prior research. A small number of prior studies have identified the use of small business credit scoring, asset-based lending, and factoring, but with some effort, more technologies may be identified. A potentially more fruitful long-run future solution would be to structure new and periodically-updated data sets to directly identify more of the lending technologies.

References

Altman, E.I., Sabato, G., 2005. Effects of the new Basel Capital Accord on bank capital requirements for SMEs. Journal of Financial Services Research 28, 15–42.

Avery, R.B., Samolyk, K.A., 2004. Bank consolidation and the provision of banking services: Small commercial loans. Journal of Financial Services Research 25, 291–325.

Bakker, Marie H.R., Klapper, L., Udell, G.F., 2004. Financing small and medium-size enterprises with factoring: Global growth and its potential in Eastern Europe. World Bank Monograph, Washington, DC. Available from: <http://wbln0018.worldbank.org/html/FinancialSectorWeb.nsf/(attachmentweb)/3342/$FILE/3342.pdf>.

Barth, J.R., Caprio Jr., G., Levine, R., 2004. Bank supervision and regulation: What works best? Journal of Financial Intermediation 13, 205–248.

Beck, T., Demirguc-Kunt, A., Maksimovic, V., 2004a. Bank competition and access to finance: International evidence. Journal of Money, Credit, and Banking 36, 627–648.

Beck, T., Demirguc-Kunt, A., Maksimovic, V., 2004b. Financing patterns around the world: Are small firms different? World Bank, working paper.

Beck, T., Demirguc-Kunt, A., Maksimovic, V., 2005. Financial and legal constraints to firm growth: Does firm size matter? Journal of Finance 60, 137–177.

Berger, A.N., 2006. Potential competitive effects of Basel II on banks in SME credit markets in the United States. Journal of Financial Services Research 29, 5–36.

Berger, A.N., Frame, W.S., 2006. Small business credit scoring and credit availability. Journal of Small Business Management 46.

Berger, A.N., Udell, G.F., 1994. Did risk-based capital allocate bank credit and cause a 'credit crunch' in the U.S.? Journal of Money, Credit and Banking 26, 585–628.

Berger, A.N., Udell, G.F., 1995. Relationship lending and lines of credit in small firm finance. Journal of Business 68, 351–382.

Berger, A.N., Udell, G.F., 2002. Small business credit availability and relationship lending: The importance of bank organisational structure. Economic Journal 112, F32–F53.

Berger, A.N., Saunders, A., Scalise, J.M., Udell, G.F., 1998. The effects of bank mergers and acquisitions on small business lending. Journal of Financial Economics 50, 187–229.

Berger, A.N., DeYoung, R., Genay, H., Udell, G.F., 2000. The globalization of financial institutions: Evidence from cross-border banking performance. Brookings–Wharton Papers on Financial Services 3, 23–158.

Berger, A.N., Goldberg, L.G., White, L.J., 2001a. The effects of dynamic changes in bank competition on the supply of small business credit. European Finance Review 5, 115–139.

Berger, A.N., Klapper, L.F., Udell, G.F., 2001b. The ability of banks to lend to informationally opaque small businesses. Journal of Banking and Finance 25, 2127–2167.

Berger, A.N., Bonime, S.D., Goldberg, L.G., White, L.J., 2004a. The dynamics of market entry: The Effects of mergers and acquisitions on entry in the banking industry. Journal of Business 77, 797–834.

Berger, A.N., Hasan, I., Klapper, L.F., 2004b. Further evidence on the link between finance and growth: An international analysis of community banking and economic performance. Journal of Financial Services Research 25, 169–202.

Berger, A.N., Clarke, G.R.G., Cull, R., Klapper, L., Udell, G.F., 2005a. Corporate governance and bank performance: A joint analysis of the static, selection, and dynamic effects of domestic, foreign, and state ownership. Journal of Banking and Finance 29, 2179–2221.

Berger, A.N., Frame, W.S., Miller, N.H., 2005b. Credit scoring and the availability, price, and risk of small business credit. Journal of Money, Credit, and Banking 37, 191–222.

Berger, A.N., Miller, N.H., Petersen, M.A., Rajan, R.G., Stein, J.C., 2005c. Does function follow organizational form? Evidence from the lending practices of large and small banks. Journal of Financial Economics 76, 237–269.

Berger, A.N., Rosen, R.J., Udell, G.F., 2006. Does market size structure affect competition? The case of small business lending. Journal of Banking and Finance, in press, doi:10.1016/j.jbankfin.2005.10.010.

Berkowitz, J., White, M.J., 2004. Bankruptcy and small firms' access to credit. The Rand Journal of Economics 35, 69–84.

Berlin, M., Loeys, J., 1988. Bond covenants and delegated monitoring. Journal of Finance 43, 397–412.

Bonaccorsi di Patti, E., Dell'Ariccia, G., 2004. Bank competition and firm creation. Journal of Money, Credit, and Banking 36, 225–251.

Boot, A., Thakor, A., 2000. Can relationship banking survive competition? Journal of Finance 55, 679–713.

Buch, C.M., 2003. Information versus regulation: What drives the international activities of commercial banks? Journal of Money Credit and Banking 35, 851–869.

Burkart, M., Ellingsen, T., 2004. In-kind finance. A theory of trade credit. American Economic Review 94, 569–590.

Burkart, M., Ellingsen, T., Giannetti, M., 2005. What you sell is what you lend? Explaining trade credit contracts. Stockholm School of Economics, Working Paper.

Carbó-Valverde, S., Rodríguez-Fernández, F., Udell, G., 2006. Bank market power and SME financing constraints. University of Granada, Working Paper.

Carey, M., Post, M., Sharpe, S.A., 1998. Does corporate lending by banks and finance companies differ? Evidence on specialization in private debt contracting. Journal of Finance 53, 845–878.

Cetorelli, N., 2004. Bank concentration and competition in Europe. Journal of Money, Credit, and Banking 36, 543–558.

Cetorelli, N., Gambera, M., 2001. Banking market structure, financial dependence and growth: International evidence from industry data. Journal of Finance 56, 617–648.

Cetorelli, N., Strahan, P.E., 2006. Finance as a barrier to entry: Bank competition and industry structure in local U.S. markets. Journal of Finance 61.

Chan, Y.-S., Kanatas, G., 1985. Asymmetric valuation and the role of collateral in loan agreements. Journal of Money, Credit and Banking 17, 85–95.

Chan, K., Chan, L.K.C., Jegadeesh, N., Lakonishok, J., 2006. Earnings quality and stock returns. Journal of Business 79.

Chemmanur, T., Yan, A., 2000. Equilibrium leasing contracts under double-sided asymmetric information. Boston College, Working Paper.

Claessens, S., Demirguc-Kunt, A., Huizinga, H., 2001. How does foreign entry affect the domestic banking market? Journal of Banking and Finance 25, 891–911.

Clarke, G., Cull, R., Martinez Peria, M.S., Sanchez, S.M., 2005a. Bank lending to small businesses in Latin America: Does bank origin matter? Journal of Money, Credit, and Banking 37, 83–118.
Clarke, G., Cull, R., Megginson, W.L., 2005b. Bank privatization. Special Issue of Journal of Banking and Finance 29.
Cole, S., 2004. Fixing market failures or fixing elections? Agricultural credit in India, Mimeo, MIT.
Cole, R.A., Goldberg, L.G., White, L.J., 2004. Cookie-cutter versus character: The micro structure of small business lending by large and small banks. Journal of Financial and Quantitative Analysis 39, 227–251.
Commercial Finance Association, 2003. Marketing Survey for the Asset-Based Financial Services Industry.
Cook, L., 1999. Trade credit and bank finance: Financing small firms in Russia. Journal of Business Venturing 14, 493–518.
Cunat, V.M., 2003. Trade credit: Suppliers as debt collectors and insurance providers. Universitat Pampeu Fabra, Working Paper.
Degryse, H., Cayseele, P.V., 2000. Relationship lending within a bank-based system: Evidence from European small business data. Journal of Financial Intermediation 9, 90–109.
Demirguc-Kunt, A., Maksimovic, V., 2002. Firms as financial intermediaries: Evidence from trade credit data, University of Maryland, Working Paper.
DeYoung, R., Nolle, D.E., 1996. Foreign-owned banks in the U.S.: Earning market share or buying it? Journal of Money, Credit, and Banking 28, 622–636.
Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2003. Courts. Quarterly Journal of Economics 118, 453–516.
European Bank for Reconstruction and Development (EBRD), 2003. Transition Report 2003.
Elsas, R., 2005. Empirical determinants of relationship lending. Journal of Financial Intermediation 14, 32–57.
Ergungor, O.E., 2004. Market- vs. bank-based financial systems: Do rights and regulations really matter? Journal of Banking and Finance 28, 2869–2887.
Fisman, R., Love, I., 2003. Trade credit: Financial intermediary development, and industry growth. Journal of Finance 58, 353–374.
Frame, W.S., Padhi, M., Woosley, L., 2004. Credit scoring and the availability of small business credit in low- and moderate income areas. Financial Review 39, 34–54.
Frank, M.Z., Maksimovic, V., 2003. Trade credit, collateral and adverse selection, University of Maryland, Working Paper.
Giannetti, M., Ongena, S., 2005. Financial integration and entrepreneurial activity: Evidence from foreign bank entry in emerging markets. European Corporate Governance Institute, Working Paper 91/2005.
Gilligan, T.W., 2004. Lemons and leases in the used business aircraft market. Journal of Political Economy 112, 1157–1180.
Goldberg, L.G., Saunders, A., 1981. The determinants of foreign banking activity in the United States. Journal of Banking and Finance 5, 17–32.
Graham, J.R., Lemmon, M.L., Shallheim, J.S., 1998. Debt, leases, taxes, and the endogeneity of corporate tax status. Journal of Finance 53, 131–162.
Guiso, L., Sapienza, P., Zingales, L., 2004. Does local financial development matter? Quarterly Journal of Economics 119, 929–969.
Hannan, T.H., 1991. Bank commercial loan markets and the role of market structure: Evidence from surveys of commercial lending. Journal of Banking and Finance 15, 133–149.
Haynes, G.W., Ou, C., Berney, R., 1999. Small business borrowing from large and small banks. In: J.L., Williams, A., Rhine, S.L.W. (Eds.), Business Access to Capital and Credit, Blanton, A Federal Reserve System Research Conference, pp. 287–327.
Hendel, I., Lizzeri, A., 2002. The role of leasing under adverse selection. Journal of Political Economy 110, 113–142.
Jappelli, T., Pagano, M., 2002. Information sharing, lending and defaults: Cross-country evidence. Journal of Banking and Finance 26, 2017–2045.
Jappelli, T., Pagano, M., Bianco, M., 2005. Courts and banks: Effects of judicial enforcement on credit markets. Journal of Money, Credit and Banking 37, 223–244.
Jayaratne, J., Wolken, J.D., 1999. How important are small banks to small business lending? New evidence from a survey of small firms. Journal of Banking and Finance 23, 427–458.
Kallberg, J., Udell, G.F., 2003. The value of private sector credit information sharing: The U.S. case. Journal of Banking and Finance 27, 449–469.
Karceski, J., Ongena, S., Smith, D., 2005. The impact of bank consolidation on commercial borrower welfare. Journal of Finance 60.

Klapper, L., 1998. Short-term collateralization: Theory and evidence, New York University, Working Paper.
Klapper, L., 2006. The role of factoring for financing small and medium enterprises. Journal of Banking and Finance, this issue, doi:10.1016/j.jbankfin.2006.05.001.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W., 1998. Law and finance. Journal of Political Economy 106, 1113–1155.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2002. Government ownership of banks. Journal of Finance 57, 265–301.
Levine, R., 2003. Denying foreign bank entry: Implications for bank interest margins. University of Minnesota, Mimeo.
Love, I., Mylenko, N., 2003. Credit reporting and financing constraints, World Bank, Working Paper.
Martinez Peria, M.S., Mody, A., 2004. How foreign participation and market concentration impact bank spreads: Evidence from Latin America. Journal of Money, Credit, and Banking 36, 511–537.
Mester, L.J., Nakamura, L.I., Renault, M., forthcoming. Checking accounts and bank monitoring, Review of Financial Studies, Working Paper.
Mian, A., forthcoming. Distance constraints: The limits of foreign lending in poor economies, Journal of Finance.
Miller, M., 2003. Credit Reporting Systems and the International Economy. MIT Press, Cambridge.
Miller, M., Rojas, D., 2004. Improving access to credit for SMEs: An empirical analysis of the viability of pooled data SME credit scoring models in Brazil, Colombia, and Mexico, World Bank, Working Paper.
Nakamura, L.I., 1993. Commercial bank information: Implications for the structure of banking. In: Klausner, M.D., White, L.J. (Eds.), Structural Change in Banking. Business One Irwin, Homewood, IL, pp. 131–160.
Padilla, A.J., Pagano, M., 1999. Sharing default information as a borrower discipline device, University of Salerno: CSEF Working Paper n. 21.
Pagano, M., Jappelli, T., 1993. Information sharing in credit markets. Journal of Finance 43, 1693–1718.
Peek, J., Rosengren, E.S., 1995. Bank regulation and the credit crunch. Journal of Banking and Finance 19, 679–692.
Petersen, M.A., Rajan, R.G., 1994. The benefits of firm-creditor relationships: Evidence from small business data. Journal of Finance 49, 3–37.
Petersen, M.A., Rajan, R.G., 1995. The effect of credit market competition on lending relationship. Quarterly Journal of Economics 110, 407–443.
Petersen, M.A., Rajan, R.G., 1997. Trade credit: Theories and evidence. Review of Financial Studies 10, 661–669.
Qian, J., Strahan, P.E., 2005. How law and institutions shape financial contracts: The case of bank loans, National Bureau of Economic Research, Working Paper 11052.
Rajan, R.G., Zingales, L., 2003. The great reversals: The politics of financial development in the twentieth century. Journal of Financial Economics 69, 5–50.
Robb, A.M., 2002. Small business financing: Differences between young and old firms. Journal of Entrepreneurial Finance and Business Ventures 7, 45–65.
Sapienza, P., 2004. The effects of government ownership on bank lending. Journal of Financial Economics 74, 357–384.
Scott, J.A., 2004. Small business and value of community financial institutions. Journal of Financial Services Research 25, 207–230.
Sharpe, S.A., 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. Journal of Finance 45, 1069–1087.
Stein, J.C., 2002. Information production and capital allocation: Decentralized vs. hierarchical firms. Journal of Finance 57, 1891–1921.
Stulz, R.M., Williamson, R., 2003. Culture, openness and finance. Journal of Financial Economics 70, 313–349.
Townsend, R.M., Yaron, J., 2001. The credit risk-contingency system of an Asian development bank, Economic Perspectives, Federal Reserve Bank of Chicago, pp. 31–48.
Udell, G.F., 2004. Asset-Based Finance. The Commercial Finance Association, New York.
Williamson, O., 1988. Corporate finance and corporate governance. Journal of Finance 43, 567–591.
World Bank, 2000. Development of Financial Leasing in Lithuania, December.
World Bank, 2002. Lithuania, Insolvency and Creditor Rights System, February.
World Bank, 2003. Czech Republic, Policy note: Insolvency and Creditor Rights System, fourth draft, October.

Journal of Financial Economics 100 (2011) 85–97

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Tests of ex ante versus ex post theories of collateral using private and public information ☆

Allen N. Berger [a,b,d,*], W. Scott Frame [c], Vasso Ioannidou [d]

[a] *Moore School of Business, University of South Carolina, 1705 College Street, Columbia, SC 29208, USA*
[b] *Wharton Financial Institutions Center, Philadelphia, PA, USA*
[c] *Federal Reserve Bank of Atlanta, USA*
[d] *CentER, Tilburg University, The Netherlands*

A R T I C L E I N F O

Article history:
Received 25 October 2009
Received in revised form
12 February 2010
Accepted 15 March 2010
Available online 9 November 2010

JEL classification:
G21
D82
G38

Keywords:
Collateral
Asymmetric Information
Banks

A B S T R A C T

Collateral is a widely used, but not well understood, debt contracting feature. Two broad strands of theoretical literature explain collateral as arising from the existence of either ex ante private information or ex post incentive problems between borrowers and lenders. However, the extant empirical literature has been unable to isolate each of these effects. This paper attempts to do so using a credit registry that is unique in that it allows the researcher to have access to some private information about borrower risk that is unobserved by the lender. The data also include public information about borrower risk, loan contract terms, and ex post performance for both secured and unsecured loans. The results suggest that the ex post theories of collateral are empirically dominant, although the ex ante theories are also valid for customers with short borrower–lender relations that are relatively unknown to the lender.

1. Introduction

Collateral is a prominent feature of debt contracts. Residential and commercial mortgages, motor vehicle and equipment loans, and interbank repurchase agreements all rely heavily on readily marketable assets to secure funding. Commercial loans only sometimes require collateral, and the pledged assets tend to be heterogeneous.

The use of collateral in debt contracts can be costly for lenders, borrowers, and (in some cases) even society at large. Lenders incur costs of screening and monitoring the pledged assets, as well as any enforcement and disposal expenses in the case of repossession (e.g., Leeth and Scott, 1989). Collateral could also impose opportunity costs on borrowers by tying up assets that might otherwise be put to more productive uses.[1] Borrowers can suffer fluctuations in

☆ We thank the following for helpful comments and suggestions: an anonymous referee, Sreedhar Bharath, Martin Brown, Marco Da Rin, Hans Degryse, Mark Flannery, Mariassunta Giannetti, Simon Kwan, Patrick Legros, Judit Montoriol-Garriga, Jan Krahnen, Bent Vale, Bas Werker, Feng Wu, conference and seminar participants at the 2010 Allied Social Sciences Association meetings, the 2010 Chicago Federal Reserve Conference on Bank Structure and Competition, the 2009 European Finance Association meetings, the 2009 Financial Intermediation Research Society meetings, the 2009 Financial Management Association meetings, the conference on "Business Models in Banking: Is There a Best Practice?," Pennsylvania State University, Sveriges Riksbank, and Tilburg University. The views expressed do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

* Corresponding author at: Moore School of Business, University of South Carolina, 1705 College Street, Columbia, SC 29208, USA. Tel.: +1 803 576 8440; fax: +1 803 777 6876.

E-mail addresses: aberger@moore.sc.edu (A.N. Berger), scott.frame@atl.frb.org (W.S. Frame), v.p.ioannidou@uvt.nl (V. Ioannidou).

[1] Renegotiations, which are frequent in private debt agreements (Smith, 1993), limit the cases in which assets are tied up in less productive uses and hence limit the opportunity costs to the borrower.

0304-405X/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jfineco.2010.10.014

their credit availability as the values of their securable assets vary, particularly for loans secured by accounts receivable or inventory or both. In certain circumstances, collateral could also result in social costs (externalities) when changes in the value of widely pledged assets, such as real estate, are correlated across borrowers. These correlated changes in collateral values could act to amplify the business cycle through procyclical changes in access to credit (e.g., Bernanke and Gertler, 1989, 1990; Kiyatoki and Moore, 1997). Recent research suggests that the significant decline in real estate collateral values in Japan in the early 1990s played an important role in reducing debt capacity and investment in that nation (Gan, 2007). A similar procyclical effect could have occurred in the US and other nations during the recent financial crisis, triggered in 2007 by the collapse in real estate collateral values.

Given that collateral is costly and yet widely employed, it is natural to inquire as to the economic functions of this contracting tool. Economic theory largely explains collateral either as an attempt to compensate for ex ante asymmetric information or as a method of reducing ex post incentive problems. Specifically, one set of theoretical models explains collateral as arising from ex ante information gaps between borrowers and lenders that can otherwise lead to an equilibrium characterized by adverse selection and credit rationing in the spirit of Stiglitz and Weiss (1981). In this case, collateral allows lenders to sort observationally equivalent loan applicants through signaling. Specifically, lenders offer a menu of contract terms such that observationally equivalent applicants with higher quality projects choose secured debt with lower risk premiums, while those with lower quality projects self-select into unsecured debt with higher risk premiums (e.g., Bester, 1985, 1987; Besanko and Thakor, 1987a, 1987b; Chan and Thakor, 1987; Boot, Thakor, and Udell, 1991). The ex ante theories predict that unobservably riskier borrowers are less likely to pledge collateral.

A second set of theoretical models motivates collateral as part of an optimal debt contract by invoking ex post frictions. These could include moral hazard concerns (e.g., Boot, Thakor, and Udell, 1991; Boot and Thakor, 1994; Aghion and Bolton, 1997; Holmstrom and Tirole, 1997); difficulties in enforcing contracts (e.g., Banerjee and Newman, 1993; Albuquerque and Hopenhayn, 2004; Cooley, Marimon, and Quadrini, 2004); and costly state verification (e.g., Townsend, 1979; Gale and Hellwig, 1985; Williamson, 1986; Boyd and Smith, 1994). The ex post theories predict that observably riskier borrowers are more likely to be required to pledge collateral.

In this paper, we test the empirical predictions generated by both the ex ante private information/signaling models and the ex post models in which collateral is used to overcome borrower/lender incentive conflicts. Our empirical test attempts to identify the effect of the two sets of theories by studying variation in the incidence of collateral pledges at loan origination. The ex ante and ex post theories are not mutually exclusive. Either or both could be consistent with the data. This test exploits differences in information that is available within a credit registry (and known to us) versus the information the registry discloses to prospective lenders. This provides clean measures of private information (information known to borrowers, but not to lenders) and public information (information known to both borrowers and lenders) with which to test the relevance of the two broad sets of collateral theories. Our results are consistent with both sets of collateral theories, although the ex ante private information theories appear to hold only for borrowers that are relatively unknown to the lender (i.e., borrowers with short relations).

The remainder of the paper is structured as follows. Section 2 provides a review of the related empirical literature. Section 3 describes the credit registry data and information-sharing regime. Section 4 outlines our empirical test and presents the results. Additional evidence is presented in Section 5 and Section 6 concludes.

2. Empirical literature review

Some of the extant empirical literature pertaining to collateral focuses on how collateral incidence relates to measures of borrower risk and proxies for private information. Consistent with the ex post theories, several studies find that observably riskier borrowers are more likely to pledge collateral. One study finds a positive association between financial leverage and collateral (Brick and Palia, 2007), another finds that firms with better public ratings are less likely to pledge collateral (Gonas, Highfield, and Mullineaux, 2004), and four others find positive associations between collateral and past observed repayment problems (Harhoff and Korting, 1998; Chakraborty and Hu, 2006; Jimenez, Salas, and Saurina, 2006; Brick and Palia, 2007).[2]

Studies testing theories concerning the importance of ex ante private information as a driver of collateral decisions have been much less successful. Many studies examine the effect of lender–borrower relation strength on collateral incidence. The idea here is that stronger relations (in terms of length, breadth, or intensity) will result in private information being revealed about the firm as lenders gather proprietary information about the borrower's character, reliability, and project choice over time (e.g., Petersen and Rajan, 1994; Berger and Udell, 1995; Degryse and Van Cayseele, 2000). However, the effect of access to private information on collateral pledges is ambiguous because this information could be favorable or unfavorable. The ex ante theories predict that unobservably safer borrowers pledge collateral.[3]

Not surprisingly, empirical tests of the role of relation strength in determining whether collateral is pledged are mixed. Berger and Udell (1995), Harhoff and Korting

[2] In a related analysis, Liberti (2009) finds that the percentage of the loan amount that is collateralized is positively related to loan risk ratings at origination for a cross-country sample of loans to small and medium-size firms.

[3] Furthermore, at least three other biases could influence the interpretation of these results. First, the more opaque the firm is, the more valuable is the private information to the lender – potentially leading to an information monopoly or "lock-in" effect that would manifest itself through positive relations (e.g., Greenbaum, Kanatas, and Venezia, 1989; Sharpe, 1990; Rajan, 1992). Second, a bias toward a positive association between collateral and relation strength could occur if lenders use relation lending technology for their most opaque borrowers (e.g., Berger, Espinosa-Vega, Frame, and Miller, 2010). Finally, the results could be biased toward a negative association to the extent that collateral and relations are substitute methods of dealing with opacity problems (e.g., Berger, Espinosa-Vega, Frame, and Miller, 2010).

(1998), Chakraborty and Hu (2006), and Brick and Palia (2007) each reports finding that stronger relations are inversely related to the incidence of collateral for loans drawn under lines of credit. Chakraborty and Hu (2006) also find that collateral is negatively related to relation scope defined as the total number of financial services used by the firm. By contrast, several studies report that the incidence of collateral is positively related to an indicator of "main bank" or "house bank" status (Machauer and Weber, 1998; Elsas and Krahnen, 2000; Degryse and Van Cayseele, 2000; Lehmann and Neuberger, 2001; Menkhoff, Neuberger, and Suwanaporn, 2006). Consistent with these latter studies, Ono and Uesugi (2009) find that relation length is positively associated with collateral for a sample of small Japanese firms.

More recent studies attempt to find other proxies for private information with which to test the ex ante theories of collateral. One study examines differences in informational opacity across borrowers, finding that more transparent firms are less likely to pledge collateral. Specifically, Gonas, Highfield, and Mullineaux (2004) find that large exchange-listed firms and those with public debt ratings are less likely to pledge collateral for bank loans. Another study exploits variation in lender information sets brought about by the use of credit scoring technology and finds that it reduces the incidence of collateral (Berger, Espinosa-Vega, Frame, and Miller, in press). While each of these studies succeeds in better identifying variation in the information environment across borrowers or lenders, none identifies borrower-specific private information as being favorable or unfavorable and, therefore, none directly tests the key empirical implication of the ex ante theories.

One recent study attempts to do this using ex post default as a measure of ex ante adverse private information, finding that this measure is negatively related to the probability of pledging collateral at origination for young firms (Jimenez, Salas, and Saurina, 2006). However, because collateral could raise the borrower's cost of default, one might expect to find that secured debt is less likely to default, irrespective of whether ex ante asymmetric information is important. Moreover, defaults could reflect moral hazard or other ex post frictions and, thus, might not isolate the effects of ex ante private information.

Our methods improve upon this literature through the availability of data that allows us to isolate the alternative theories of collateral empirically: ex ante private information versus ex post contracting frictions. Specifically, our test exploits differences in information that is available within a credit registry (and known to us) versus the information the registry discloses to prospective lenders. This provides clean measures of private and public information with which to test the relevance of the two sets of collateral theories.

3. Data and information environment

Our analysis utilizes data from the Central de Información de Riesgos Crediticios (CIRC), the public credit registry of Bolivia, provided by the Bolivian Superintendent of Banks and Financial Entities (SBEF). Since CIRC's creation in 1989, the SBEF requires all formal (licensed and regulated) financial institutions operating in Bolivia to record information on all loans. Our sample covers the entire credit registry for the period between January 1998 and December 2003. For each loan, we have information on the date of origination, maturity date, contract terms, and ex post performance through the sample period. For each borrower, we have information about its industry, physical location, legal structure, banking relations, and whether it has been delinquent or defaulted on a loan in the recent past.

The SBEF requires that some loan information is shared among the participating institutions to help alleviate the otherwise pervasive information asymmetries in the Bolivian credit markets. After written authorization from a prospective customer, a lender can access the registry and obtain a credit report, which contains information on all outstanding loans of the customer for the previous two months. Entries include originating bank, loan amount, type of loan, value of collateral, value of overdue payments, and the borrower's credit rating from the originating bank. Loans with overdue payments remain in the registry until they are paid off, even if they are past maturity. This implies that delinquencies in the past two months and past defaults from any previous period are observable to other lenders through the registry. By contrast, delinquencies that were paid off more than two months ago are not observable to other lenders through the registry (Campion, 2001).

An underlying assumption that we maintain in the paper is that at least some of the information about past delinquencies does not become observable through other sources. We think that this was a reasonable assumption for several reasons. First, during the sample period, no other credit registry is operational in Bolivia (de Janvry, Sadoulet, McIntosh, Wydick, Luoto, Gordillo, and Schuetz, 2003).[4] Second, Bolivian credit markets are opaque. The vast majority of firms do not have audited financial statements, and the quality of existing financial statements is poor as many firms engage in tax evasion (Sirtaine, Skamnelos, and Frank, 2004). Third, evidence presented in Ioannidou and Ongena (2010) is consistent with the assumption that at least some of this information remains unobserved as borrowers appear to use the two-month disclosure window strategically. The authors find that banks are unwilling to extend credit to new customers with observable repayment problems and that borrowers trying to switch to new banks clear past due payments on their outstanding loans for those

[4] Until April 2000, Bolivian law prohibited the creation of private credit registries. Once the law changed, private credit registries started to emerge, but were not operational before the end of our sample period. Alternative informal information sharing mechanisms, however, were in place among microfinance institutions whose clients were inadequately covered by the CIRC; only formal financial institutions report to the CIRC, but many microfinance institutions are informal. Hence, to obtain information about high-risk clients, many microfinance institutions maintained internal "blacklists" and shared them with each other. They also used credit registry type services offered by two NGO associations, FINRURAL and CIPAME, and purchased publicly available information (e.g., penal and civil judgments, newspaper articles, and business failures) collected by SIPROTEC and DATOS. The poor coverage and quality of these services was blamed for the over-indebtedness crisis of the microfinance sector in the late 1990s and has led to the change in the law in 2000 (see Campion, 2001; de Janvry, Sadoulet, McIntosh, Wydick, Luoto, Gordillo, and Schuetz, 2003).

two months, manage to switch, but tend to return to nonperformance soon thereafter.

By having access to the entire credit registry, this information-sharing regime allows us to construct our indicators of observed and unobserved borrower risk, our key independent variables for testing the two collateral theories. We construct three indicators of observed risk: a dummy variable that equals one if the loan is given to a borrower that defaulted with any bank in the previous 12 months (*Default_Observable_Registry*), a dummy variable that equals one if the borrower had been 30+ days delinquent with any bank in the previous two months (*Npl_Observable_Registry*), and a dummy variable that equals one if the loan is given to a borrower that had been 30+ days delinquent with the same bank anytime from three to 12 months prior (*Npl_Observable_Relation*).[5] To measure unobserved risk, we create a dummy variable that equals one if the borrower had 30+ day delinquencies at other banks three to 12 months prior to the loan origination (*Npl_Unobservable*). Our empirical results are materially unchanged if we expand the performance horizon back 18 or 24 months.

The data include loans from commercial banks and nonbank financial institutions (e.g., microfinance institutions, credit unions, mutual societies, and general deposit warehouses).[6] To keep the set of lenders and borrowers homogenous in terms of financial structure and regulation, we focus exclusively on commercial loans granted by commercial banks between March 1999 and December 2003.[7] Table 1 provides a list of the 13 commercial banks that were active in Bolivia during the sample period, seven of which were foreign-owned (four branches and three subsidiaries).[8] As shown in Table 1, five banks dominated the Bolivian banking sector during this time, each with total assets averaging at least $500 million and with more than 10% market share of deposits and loans. The Herfindahl-Hirschman index (sum of squares of market share percentages) for deposits is 1,292 and for loans it is 1,236, suggesting moderate market concentration.

Over the 1998–2003 period, the Bolivian economy slowed markedly due to the Russian and Asian financial crises, which resulted in currency depreciations for many emerging market currencies, including the Bolivian peso.[9] While Bolivian bank assets and liabilities were largely denominated in US dollars, wages and business income were largely paid in domestic currency. Hence, the currency depreciation had the effect of inflating debt-to-income ratios for consumers and firms, sending some into financial distress. As a result, bank balance sheets shrunk due to deposit withdrawals and weakened loan demand, and nonperforming loans increased.[10] Despite the weakened macroeconomic and banking environment in Bolivia during our sample period, bank capital–adequacy ratios remained above the 10% minimum requirement and actually increased over time, owing to shrinking balance sheets and government intervention in the form of low-cost loans and capital injections via subordinated debt (Sirtaine, Skamnelos, and Frank, 2004).

Several types of commercial credit contracts are in the data, including credit cards, overdrafts, installment loans, discount loans, and lines of credit. We focus exclusively on installment loans and discount loans and refer to these as standard debt contracts. These contracts account for 92% of the total value of commercial loans during the sample period. Of these contracts, 98% are denominated in US dollars, and we use only these loans in our analysis.

Our sample encompasses 32,286 bank loans made to 2,676 different firms. A loan is defined by a unique identification code (loan ID) and a date of origination. This includes not only new loans to new or existing customers, but also renegotiations of previous loans. Banks, however, are required to indicate whether a new loan is a renegotiation of a previous (performing or nonperforming) loan, and we use this information to exclude renegotiations.[11] We also do not include new loans drawn on preexisting lines of credit.[12]

Table 2 provides variable names, definitions, and summary statistics for all loans in the sample and for secured and unsecured loans separately.[13] Collateral was pledged for 24.4% of the loans in the sample. As in the US, secured

[5] Default occurs when the overdue payments are persistent enough for the bank to downgrade a loan to a rating of five and write off the overdue amount. According to regulations, small loans (with an average amount smaller than $75,000) are downgraded to five if there are overdue payments for at least 121 days for secured loans and 91 days for unsecured loans. Large loans, by contrast, are downgraded to five when the borrower is considered insolvent (i.e., when the borrower's net worth is close to zero). In the sensitivity analysis below, we investigate the robustness of our findings by reestimating our models separately for loans with contract amounts above or below $75,000.

[6] Bolivia's corporate sector is segmented. It is composed of a few thousand formal firms and a lot of very small firms that are mostly informal. The formal firms have access to commercial banks (our sample), while the informal firms can only borrow from informal financial institutions (e.g., microfinance non-governmental organizations (NGOs) and foundations).

[7] Although we have data as of January 1998, we start our sample in March 1999 because prior to this date the data do not allow us to distinguish been commercial and consumer loans. However, we use the prior information from January 1998 through February 1999 to help fill in history on loans and relations that existed as of March 1999.

[8] Foreign-owned banks operating in Bolivia have similar rights and responsibilities as domestically owned institutions. One of the foreign branches, ABN Amro, left the Bolivian market in November 2000.

[9] The average annual growth rate of real gross domestic product (GDP) in Bolivia during the sample period was 2.2%, ranging between 0.7% and 3.7%.

[10] Bolivian nonperforming bank loans increased from $172 million (4.3% of total loans) in March 1999 to a peak of $570 million (20.3% of total loans) in June 2003.

[11] To the extent that some renegotiations are not recorded (either because of reporting errors or because banks do that intentionally to reduce their loan loss reserves), our sample would include some renegotiations as new loans. Hence, in the sensitivity analysis below we try to control for this possibility by dropping all suspected renegotiations from our sample.

[12] When a borrower draws on a preexisting line of credit, a new loan appears in the registry with origination date and contract terms as of the date the bank originated the credit line. Because the date the loan first appears in the registry is subsequent to the origination date, we can identify when a new loan is a draw on a preexisting line of credit and exclude it from our sample.

[13] For relation length, loan amount, and maturity we report summary statistics for the level of these variables, but our empirical models incorporate the natural logarithm of one plus the level.

Table 1
Summary statistics for commercial banks operating in Bolivia.

This table provides summary statistics on all commercial banks that were active in Bolivia between March 1999 and December 2003. *Assets* is equal to the average value of total assets in millions of US dollars during the sample period. *Deposits Share* is equal to average ratio of bank deposits to the total deposits in the banking system. Similarly, *Loans Share* is equal to the average ratio of total bank loans to the total loans in the banking system. The *Capital Ratio* reports the average ratio of total capital (Tier 1+Tier 2) to total risk-weighted assets. The *NPL Ratio* is equal to each bank's average ratio of nonperforming loans (delinquent of at least 30 days) to total loans. *Ownership* indicates whether a bank is foreign-owned or domestically owned and for foreign-owned whether it is a branch (B) or subsidiary (S). Banks for which at least 50% of their equity is foreign-owned are defined as foreign.

Bank name	*Assets*	*Deposits Share*	*Loans Share*	*Capital Ratio*	*NPL Ratio*	*Ownership*
Banco Santa Cruz	859.138	0.183	0.161	18.276	0.168	Foreign (S)
Banco Industrial	677.694	0.127	0.151	12.504	0.097	Domestic
Banco Nacional de Bolivia	621.061	0.149	0.139	11.343	0.110	Domestic
Banco Mercantil	598.541	0.142	0.125	12.076	0.091	Domestic
Banco de Crédito de Bolivia	591.024	0.134	0.126	13.985	0.130	Foreign (S)
Banco de la Unión	450.655	0.088	0.104	12.479	0.166	Domestic
Banco Económico	287.374	0.062	0.067	15.074	0.099	Domestic
Citibank	265.291	0.044	0.047	18.835	0.312	Foreign (B)
Banco Ganadero	205.477	0.042	0.046	11.888	0.105	Domestic
Banco Solidario	95.932	0.019	0.024	18.346	0.103	Foreign (S)
Banco do Brasil	31.771	0.005	0.003	54.374	0.071	Foreign (B)
Banco de la Nación Argentina	28.649	0.004	0.006	36.476	0.290	Foreign (B)
ABN Amro	22.341	0.003	0.003	42.520	0.050	Foreign (B)

debt in Bolivia has effective priority over unsecured debt in bankruptcy (see Djankov, McLiesh, and Shleifer, 2007).[14] Of the 29,485 loans that matured before the end of the sample period, 6.4% had ex post delinquencies or defaults, with secured loans having a substantially higher incidence of repayment problems (9.1% as opposed to 5.5% for unsecured loans), suggesting that banks require collateral from riskier borrowers. The data also indicate that only 0.3% of the sample loans were given to borrowers that had defaulted in the prior 12 months (*Default_Observable_Registry*). This suggests that borrowers that default rarely get another loan, either because they are credit rationed or cease to exist as a going concern. Some 5.7% of the loans were issued to firms that had been delinquent with any bank in the two prior months (*Npl_Observable_Registry*). The data also show that 7.4% of the credits were issued to firms that had been delinquent with the same bank anytime from three to 12 months prior (*Npl_Observable_Relation*). Finally, 14.0% of loans were given to borrowers with delinquencies at other banks from three to 12 months prior to the loan origination (*Npl_Unobservable*). This last information item is not revealed to the lender through the credit registry but, as our empirical analysis below suggests, it might be revealed through lending relations.

Turning to the control variables, the average banking relation in the sample is 23.1 months. This is defined as the number of months since the first loan of a particular borrower from a particular bank in the registry since January 1998. Most of the sample firms are corporations (71.4%). Partnerships (14.0%) and sole proprietorships (12.5%) are much less common. Almost one-half of the sample is composed of installment loans. The average loan amount is $161,490 and the average loan maturity is about 12 months. As expected, secured and unsecured loans have different contract terms. On average, secured loans are more than twice as large and have maturities that are six months longer.

4. Empirical analysis

Our empirical test relates the incidence of collateral to measures of observed and unobserved borrower risk, the length of the banking relation, loan- and firm-level control variables, and bank and time fixed effects. This model, which is estimated using probit, can be summarized as

$$P(Collateral_{ijt}) = f(Observed\_Risk_{ijt},\ Unobserved\_Risk_{ijt},\ Unobserved\_Risk_{ijt} *Rel\_Length_{ijt},\ Rel\_Length_{ijt},\ Firm_{ijt},\ Loan_{ijt},\ \alpha_j,\ \gamma_t), \quad (1)$$

where $P(\cdot)$ indicates probability, $Collateral_{ijt}$ is a dummy variable that equals one if the loan is secured, and i, j, and t index loans, banks, and time, respectively. The key exogenous variables are those capturing observed and unobserved borrower risk as defined in Table 2. Specifically, $Observed\_Risk_{ijt}$ is composed of the three dummy variables that indicate the observed riskiness of the borrower at loan origination: *Default_Observable_Registry*, *Npl_Observable_Registry*, and *Npl_Observable_Relation*. $Unobserved\_Risk_{ijt}$ is composed of a single dummy variable, *Npl_Unobservable*. This variable is in the credit registry and thus available to us, but not available to the lender. By using these four indicators of borrower risk, we assume that past performance, observable or unobservable, is predictive of future performance. That is, we assume that borrowers with past repayment problems are more likely to have delinquencies or defaults on future loans. The Appendix demonstrates that this is the case.

A positive, statistically and economically significant coefficient for any of the three variables included in $Observed\_Risk_{ijt}$ would be evidence in favor of the ex post theories. That is, observably risky borrowers are more likely to post collateral. By contrast, a negative, statistically and economically significant coefficient on $Unobserved\_Risk_{ijt}$ would be consistent with the ex ante theories. That is,

[14] The data used in Djankov, McLiesh, and Shleifer (2007) are available at http://www.economics.harvard.edu/faculty/shleifer/dataset.

Table 2
Variables and summary statistics.

The table reports the notation and definitions of variables used in the analysis as well as summary statistics for all loans and for secured and unsecured loans separately. With the exception of the summary statistics for the variable *Ex_Post_Nonperformance*, the number of observations is 32,286 for all loans, 7,864 for secured loans, and 24,422 for unsecured loans. For *Ex_Post_Nonperformance*, the summary statistics use the number of loans that matured before the end of the sample period: 29,485 for all loans, 7,106 for secured loans, and 22,379 for unsecured loans.

Variables	Description	All		Secured		Unsecured	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Dependent variables							
Collateral	Equals one if collateral was pledged at loan origination and zero otherwise	0.244	0.429	1.000	0.000	0.000	0.000
Risk Premium	Interest rate minus the six-month US Treasury bill rate	9.763	2.563	9.410	2.962	9.877	2.409
Ex_Post_Nonperformance	Equals one if a loan is 30+ days overdue anytime after origination or if it is downgraded to the default status (a rating of 5) and zero otherwise	0.064	0.244	0.091	0.288	0.055	0.227
Independent variables							
Observed risk							
Default_Observable_Registry	Equals one if the borrower had defaulted on a loan anytime in the previous 12 months with any lender and zero otherwise	0.003	0.052	0.004	0.067	0.002	0.046
Npl_Observable_Registry	Equals one if the borrower had overdue payments of at least 30 days with any bank anytime from $t-1$ to $t-2$ and zero otherwise	0.057	0.231	0.073	0.260	0.051	0.221
Npl_Observable_Relation	Equals one if the borrower had overdue payments of at least 30 days with the current bank anytime from $t-3$ to $t-12$, and zero otherwise	0.074	0.261	0.069	0.253	0.075	0.264
Unobserved risk							
Npl_Unobservable	Equals one if the borrower had overdue payments of at least 30 days with another bank anytime from $t-3$ to $t-12$, and zero otherwise	0.140	0.347	0.158	0.365	0.134	0.340
Relation characteristic							
Rel_Length	Length of bank–firm relation in months	23.102	16.046	22.910	16.797	23.164	15.797
Firm characteristics							
Sole Proprietorship	Equals one if the firm is a sole proprietorship and zero otherwise	0.125	0.331	0.131	0.338	0.124	0.329
Partnership	Equals one if the firm is a partnership (i.e., all or some partners have unlimited liability), and is zero otherwise	0.140	0.347	0.136	0.343	0.141	0.348
Corporation	Equals one if the firms is a corporation (i.e., all or some partners have limited liability), and is zero otherwise	0.714	0.452	0.707	0.455	0.716	0.451
Other	Equals one if the firm is a public company, a municipality, or a cultural, sport, or religious association and is zero otherwise	0.020	0.142	0.026	0.158	0.019	0.136
Loan characteristics							
Installment	Equals one if an installment loan and zero if a discount loan	0.471	0.499	0.498	0.500	0.462	0.499
Loan Amount	Loan amount at loan origination in US dollars	161,490	467,960	285,766	754,860	121,472	315,376
Maturity	Number of months between loan origination and maturity	11.880	16.308	16.444	24.162	10.411	12.440
Interest Rate	Annual contractual interest rate at loan origination	13.449	2.886	12.783	3.127	13.664	2.770
Fixed effects							
Industry	Set of dummy variables controlling for the firm's industry; there are 19 industry categories: Agriculture and cattle; Farming; Forestry and fishery; Extraction of oil and gas; Minerals; Manufacturing; Electricity, gas, and water; Construction; Wholesale and retail trade; Hotels and restaurants; Transport, storage, and communications; Financial Intermediation; Real estate activities; Public administration defense, and compulsory social security; Education; Communal and personal social services; Activities of households as employees of domestic personnel; Activities of extraterritorial organizations and bodies; and Other						
Region	Set of dummy variables controlling for region of loan origination (Chuquisaca, La Paz, Cochabamba, Oruro, Potosi, Tarija, Santa Cruz, Beni, Pando, US, Argentina, Paraguay, and Panama)						
Bank	Set of dummy variables controlling for the bank that originated the loan; there are 13 banks						
Time	Set of dummy variables controlling for the time of loan origination; there are 57 months from March 1999 to December 2003						

according to models of signaling, firms with private information that they are good are more likely to pledge collateral. Such a finding would require that the $Unobserved\_Risk_{ijt}$ variable does not indirectly become observable through other sources. For example, some of this information might have become known to the lender if the borrower previously applied for a loan to the same bank that was not granted or if this information is demanded in loan applications. To the extent that this occurs, it would bias our results against finding evidence consistent with the ex ante theories. In fact, in the limiting case in which all of the information in $Unobserved\_Risk_{ijt}$ becomes observable to the originating bank, we should find that the estimated coefficient of $Unobserved\_Risk_{ijt}$ is positive and statistically significant.

To account for the likelihood that a firm's private information declines in the length of the bank–firm relation, we also include the interaction term $Unobserved\_Risk_{ijt}*Rel\_Length_{ijt}$, where Rel_Length_{ijt} is measured as the natural logarithm of one plus the number of months that we observe the bank and borrower in a relation. We expect that the empirical relevance of the ex ante theories diminishes as the length of a bank–firm relation increases, suggesting a negative coefficient for $Unobserved\_Risk_{ijt}$ and a positive coefficient for $Unobserved\_Risk_{ijt}*Rel\_Length_{ijt}$. In other words, borrowers with favorable private information choose to pledge collateral to signal their quality only when relations are short and the bank does not know their quality. Finally, the expected sign of coefficient of Rel_Length_{ijt} by itself is ambiguous because, as argued above, the private information revealed through a relation could be favorable or unfavorable.

The vector $Firm_{ijt}$ accounts for differences in firm characteristics such as legal structure, industry, and region. We use a set of dummy variables indicating the legal structure of the firm: *Partnership*, *Corporation*, and *Other* (*Sole_Proprietorship* is the omitted group). *Industry* is a set of dummy variables controlling for the firm's industry classification (such as the Standard Industrial Classification or North American Industry Classification System codes). *Region* is a set of dummy variables that indicate the location from which the bank originated the loan. This includes nine regions in Bolivia as well as Argentina, Paraguay, Panama, and the United States.

The vector $Loan_{ijt}$ accounts for differences in the individual loan contract terms. However, each of these terms could be determined simultaneously with collateral and is potentially endogenous. We therefore estimate our empirical models both with and without these variables. *Installment* is a dummy variable equal to one if the contract is an installment loan, not a discount loan. *Loan_Amount* is measured as the natural logarithm of one plus the amount of the loan proceeds at origination in US dollars. *Maturity* is the natural logarithm of one plus the number of months between loan origination and maturity. We explicitly exclude the loan interest rate because it is jointly determined with collateral under the ex ante theories.

Bank and time (month) fixed effects are also included in the model, represented by the scalars α_j and γ_t, respectively. Bank fixed effects should capture any systematic differences in bank propensities to require collateral for their commercial loans. The time fixed effects are intended to account for temporal differences in required collateral related to the business or credit cycle. Accounting for such variation could be important given the volatile macro-financial environment in Bolivia during the period under study.

Estimation results are presented in Table 3. In Column 1, we report a benchmark specification without interaction terms and loan characteristics. In Column 2, we include the interaction term between the length of a bank–firm relation and unobserved risk, and in Column 3, we add loan characteristics. Under the heading "Probit coefficients," we report the estimated coefficients of the three probit specifications. Under the heading "Marginal effects," we report the change in probability of pledging collateral for each one of the independent variables, holding all other independent variables at their sample means. For dummy variables, we report the estimated effect of a change from zero to one.

In all three specifications, our three indicators of observed risk are positively associated with the incidence of collateral. These findings are consistent with the ex post theories, under which observably riskier borrowers are asked to pledge collateral to mitigate frictions such as moral hazard. Each of these indicators of previous delinquencies or defaults is estimated to be associated with a 3.9–12.8 percentage point estimated increase in the probability of collateral being pledged. These findings are economically significant, given that the predicted probability of collateral at the mean of all independent variables (P0) is only about 20%.

The estimated coefficient of *Unobserved Risk* is neither statistically nor economically significant in Column 1. However, when we include the interaction term between unobserved risk and relation length, the measured effect of unobserved risk becomes both statistically and economically significant. For new borrowers, for whom relation length is zero, unobserved risk is associated with 13.7–17.0 percentage point decrease in the probability of collateral, consistent with the ex ante theories. Combining the marginal effects of Unobserved Risk and the interaction term suggests that the effect is negative for relations under approximately seven months.[15] This is consistent with a reduction in private information with longer relations, as predicted by the relation lending literature.[16] Thus, the data suggest that the ex ante theories hold only for loans to borrowers with relatively short relations when asymmetric information problems are more likely to be present. Relation length itself is significantly negatively related to the incidence of collateral.

The incidence of collateral is lower for loans to partnerships or corporations than for those to sole proprietorships (the omitted category), consistent with collateral being more likely for opaque firms. Installment loans are less likely to have collateral pledged, but larger loans and those with longer maturities are more likely to be associated with collateral pledges.

[15] In Column 2, the effect of unobserved risk equals zero for relation length x when $-0.17+0.082*\ln(1+x)=0$, which is solved for x equal to 6.9. For the estimates in Column 3, the corresponding x equals 6.7. This result that significant relation information is transmitted only for a limited time is consistent with Cole (1998), who finds that most of the borrower information is collected within the first year of a relation.

[16] One way in which private information might be revealed to the bank when relations are longer is through past draws on the credit registry.

Table 3
Determinants of collateral.

This table reports probit regressions for *Collateral*, a dummy variable that equals one if the loan is secured and is equal to zero otherwise. Under the heading "Probit coefficients," we report the estimated coefficients of the three probit specifications. Standard errors, corrected for heteroskedasticity, are reported between brackets. Under the heading "Marginal effects," we report the change in probability of pledging collateral for each one of the independent variables. For continuous variables we report the effect for an infinitesimal change in each independent variable, and for dummy variables we report the estimated effect of a change from zero to one. P0 is the predicted probability that collateral is pledged evaluated at the mean of all independent variables. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variables	Probit coefficients			Marginal effects		
	(1)	(2)	(3)	(1)	(2)	(3)
Observed risk						
Default_Observable_Registry	0.388**	0.369**	0.335**	0.128**	0.12**	0.104**
	[0.160]	[0.161]	[0.161]	[0.058]	[0.058]	[0.056]
Npl_Observable_Registry	0.222***	0.219***	0.262***	0.069***	0.068***	0.079***
	[0.040]	[0.040]	[0.040]	[0.013]	[0.013]	[0.013]
Npl_Observable_Relation	0.163***	0.144***	0.136***	0.05***	0.043***	0.039***
	[0.035]	[0.036]	[0.037]	[0.011]	[0.011]	[0.011]
Unobserved risk						
Npl_Unobservable	0.043	−0.765***	−0.622***	0.012	−0.17***	−0.137***
	[0.027]	[0.094]	[0.094]	[0.008]	[0.015]	[0.016]
*Npl_Unobservable*Rel_Length*		0.287***	0.243***		0.082***	0.067***
		[0.031]	[0.031]		[0.009]	[0.009]
Relation characteristic						
Rel_Length	−0.148***	−0.164***	−0.131***	−0.043***	−0.047***	−0.036***
	[0.010]	[0.010]	[0.010]	[0.003]	[0.003]	[0.003]
Firm characteristics						
Partnerships	−0.211***	−0.214***	−0.267***	−0.057***	−0.057***	−0.067***
	[0.037]	[0.037]	[0.038]	[0.009]	[0.009]	[0.009]
Corporations	−0.074***	−0.078***	−0.153***	−0.022***	−0.023***	−0.043***
	[0.027]	[0.027]	[0.028]	[0.008]	[0.008]	[0.008]
Other	0.164**	0.154**	−0.021	0.05**	0.047**	−0.006
	[0.065]	[0.065]	[0.068]	[0.021]	[0.021]	[0.018]
Loan characteristics						
Installment Loan			−0.14***			−0.038***
			[0.025]			[0.007]
Loan Amount			0.141***			0.039***
			[0.007]			[0.002]
Maturity			0.372***			0.102***
			[0.015]			[0.004]
Industry, *Region*, *Bank*, and *Time* dummy variables	Included	Included	Included	Included	Included	Included
P0				0.209	0.208	0.193
Pseudo *R*-squared	0.213	0.215	0.264	0.213	0.209	0.264
Number of observations	32,286	32,286	32,286	32,286	32,286	32,286

We conduct four robustness checks. First, we include a number of additional relation characteristics to the specification reported in Column 2 of Table 3. Specifically, we include a dummy variable for multiple banking relations, a dummy variable for when the bank is the firm's primary lender (i.e., more than 50% of outstanding loan balances are from that bank), and a dummy variable for the existence of other lending products such as other loans, credit cards, credit lines, and overdrafts in the current account. The empirical results are materially unchanged by including additional relation metrics.

Second, we estimate the model separately for installment and discount loans. The results are very similar to those reported in Column 2 of Table 3. One difference is that the estimated coefficients of *Default_Observable_Registry* and *Npl_Observable_Relation* in the discount loans equations are not statistically significant. However, the estimated coefficient of the third indicator of observed borrower risk, *Npl_Observable_Registry*, remains positive and statistically significant, supporting the ex post theories.

Third, we estimate the model separately for loans with contract amounts above and below $75,000. The results are very similar to those reported earlier, with the exception of the coefficient of *Default_Observable_Registry* that is not statistically significant in the sample of loans with a contract amount below $75,000.

Finally, we include the loan interest rate at origination to the vector of loan characteristics included in the specification reported in Column 3 of Table 3. While the interest rate is found to have negative and statistically significant relations with the incidence of collateral, including this variable had no material effect on our results.

5. Additional evidence

We believe that studying the effects of observed and unobserved borrower risk on the incidence of collateral is the most appropriate way for empirically testing the two sets of collateral theories because it allows identification of the individual effects. However, an empirical literature relates measures of borrower risk (such as the loan risk premium or ex post loan performance) to whether or not collateral was pledged for a given credit. In this case, one might be able to surmise whether the ex ante private information theories or the ex post incentive conflict models dominate empirically, but only under certain circumstances.

Most of the studies relating borrower risk to the incidence of collateral use the risk premium paid on the credit (yield less the risk-free rate) as the borrower risk measure. Several studies report positive relations between risk premiums and collateral pledges (Berger and Udell, 1990; Blackwell and Winters, 1997; Machauer and Weber, 1998; John, Lynch, and Puri, 2003; Brick and Palia, 2007). Two other studies report negative relations (Degryse and Van Cayseele, 2000; Lehmann and Neuberger, 2001). Jimenez and Saurina (2004) use ex post loan nonperformance as measure of borrower risk and find that loan defaults are positively related to collateral pledges.[17]

For comparative purposes, we estimate similar empirical relations using loan risk premiums and ex post loan nonperformance as risk measures. We define $Risk\_Premium_{ijt}$ as the loan interest rate (at origination) minus the rate on the six-month US Treasury bill at the end of the same month.[18] $Ex\_Post\_Nonperformance_{ijt}$ is a dummy variable that equals one if the loan eventually becomes delinquent or defaults.[19] Both of these measures are regressed on a dummy variable indicating that collateral was pledged, the length of the banking relation, the interaction of these two variables, loan- and firm-level control variables, and bank and time fixed effects summarized as

$$\begin{aligned} &Risk\_Premium_{ijt} \\ &\quad = g(Collateral_{ijt},\ Collateral_{ijt}*Rel\_Length_{ijt}, \\ &\qquad Rel\_Length_{ijt},\ Firm_{ijt},\ Loan_{ijt},\ \alpha_j,\ \gamma_t) \end{aligned} \tag{2}$$

and

$$\begin{aligned} &P(Ex\_Post\_Nonperformance_{ijt}) \\ &\quad = h(Collateral_{ijt},\ Collateral_{ijt}*Rel\_Length_{ijt}, \\ &\qquad Rel\_Length_{ijt},\ Firm_{ijt},\ Loan_{ij},\ \alpha_j,\ \gamma_t). \end{aligned} \tag{3}$$

From a theoretical perspective, the measured effect of collateral on borrower risk in Eqs. (2) and (3) is ambiguous. Under the ex post theories, collateral is required of observably riskier borrowers who are more likely to pay higher interest rates and have performance problems. But this effect could be offset, or even overturned, to the extent that collateral mitigates or eliminates the ex post frictions. The effective priority of secured debt in bankruptcy mitigates incentives for moral hazard (e.g., it reduces incentives to substitute into riskier assets or projects and prevents borrowers from obtaining additional debt that jeopardizes the lender's claims) and reduces the lender's loss given default (e.g., by increasing the seniority of secured debt over unsecured debt and by facilitating the repossession of the property and thus reducing foreclosure costs). Hence, all else equal, the effective priority of secured debt should lead to smaller loan rate premiums (Smith and Warner, 1979a, 1979b).

Under the ex ante private information theories, the measured effect of collateral is expected to be negative because it is the unobservably safer borrowers who pledge collateral more often and hence pay lower interest rates and are less likely to have performance problems. Thus, a positive measured effect of collateral on risk premiums or ex post nonperformance would suggest a net empirical dominance of the ex post theories. By contrast, a negative measured effect would suggest either a net empirical dominance of the ex ante private information theories, an overcompensating effect of collateral under the ex post theories, or both.

As in Eq. (1), we also include the interaction term between collateral and relation length, $Collateral_{ijt}*Rel\_Length_{ijt}$, because the ex ante theories are less likely to hold when the relation is long and the bank has had time to discover more of the private information about the firm. Thus, we expect a positive sign on the interaction term in Eqs. (2) and (3), as the ex post theories are more likely to empirically dominate when relations are longer. All firm and loan control variables and bank and time fixed effects are the same as in Eq. (1).

Ordinary least squares (OLS) estimates for Eq. (2) are presented in Table 4. Column 1 reports a benchmark specification without the collateral–relation length interaction term or loan characteristics. Column 2 introduces the interaction term, and Column 3 also includes the potentially endogenous loan characteristics. In all three specifications, we find strong negative associations between loan risk premiums and the incidence of collateral. This suggests that borrowers pledging collateral generally receive a lower interest rate, consistent with either a net empirical dominance of the ex ante private information theories, an overcompensating effect of collateral under the ex post theories, or both. However, it is also the case that this negative effect is mitigated to some extent by long-term borrower–lender relations, suggesting that as private information is revealed over time the discounts dissipate (see Columns 2 and 3). Consistent with our previous findings, this result suggests that the ex post theories are more relevant when relations are more established.

With respect to the control variables, we find that relation length has a positive and statistically significant coefficient in Column 1 when an interaction term is not included, while it is essentially zero in Columns 2 and 3. We also find that firms organized as partnerships and corporations pay lower average loan risk premiums than sole

[17] In an earlier study, Berger and Udell (1990) use bank-level data to study the association between the quality of the loan portfolio and the proportion of loans that were collateralized, and find evidence that collateralized loans are riskier.

[18] The six-month US Treasury rate is used because the median loan in the sample has a maturity of seven months and all loans are denominated in US dollars. Estimating relations using the rate on the three-month Treasury bill has no material effect on the results.

[19] For consistency with Jimenez and Saurina (2004), we also estimate relations using the probability of default as the dependent variable in place of the probability of delinquency or default (i.e., we adopt a more conservative definition of nonperformance). Again, the signs and significance of the results are virtually unchanged.

Table 4
Determinants of loan risk premiums.

This table reports ordinary least squares (OLS) regressions for *Risk_Premium$_{ijt}$*, which is defined at the loan interest rate less the six-month US Treasury bill rate at the end of the same month. Standard errors, corrected for heteroskedasticity, are reported between brackets. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variables	OLS coefficients		
	(1)	(2)	(3)
Loan characteristics			
Collateral	−0.419***	−0.664***	−0.292***
	[0.034]	[0.085]	[0.079]
*Collateral*Rel_Length*		0.09***	0.076***
		[0.029]	[0.027]
Relation characteristic			
Rel_Length	0.038**	0.007	0.001
	[0.015]	[0.017]	[0.015]
Firm characteristics			
Partnerships	−0.245***	−0.245***	−0.103**
	[0.046]	[0.046]	[0.043]
Corporations	−0.666***	−0.666***	−0.382***
	[0.035]	[0.035]	[0.032]
Other	−0.384***	−0.372***	0.161**
	[0.078]	[0.077]	[0.073]
Other loan characteristics			
Installment Loan			0.651***
			[0.030]
Loan Amount			−0.528***
			[0.009]
Maturity			−0.231***
			[0.018]
Industry, *Region*, *Bank*, and *Time* dummy variables	Included	Included	Included
R-squared	0.38	0.38	0.46
Number of observations	32,286	32,286	32,286

proprietorships. This is consistent with proprietorships generally being riskier and more informationally opaque than other types of firms. Regarding contract terms, it appears that installment loans carry higher interest rates, and larger loans and loans with longer maturities carry lower interest rates.

Table 5 presents the results for *Ex_Post_Nonperformance$_{ijt}$*, both in terms of the probit coefficients and marginal effects. For this analysis, we drop all loans that do not mature before the end of the sample (December 2003), thereby leaving 29,485 bank loans. Because this has the effect of reducing the average loan maturity in our sample, we also eliminate all loans originated during the last six months of the sample (July–December 2003), further reducing the sample to 28,758 loans.

In Column 1, collateral is positively associated with ex post delinquencies or defaults, consistent with net empirical dominance of the ex post theories. The estimated marginal effect suggests a 4.1 percentage point increase in the probability of ex post nonperformance for secured loans. This effect is economically significant, because the predicted probability of ex post nonperformance at the mean of all independent variables (P0) is 4.7%. This suggests that secured loans are almost twice as likely to have repayment problems as unsecured loans. However, when the interaction term is introduced in Column 2, the results change substantially. The coefficient on collateral becomes zero, suggesting that the net effect of collateral when the customer is new to the bank (i.e., when relation length is zero) is zero. However, the positive coefficient of the interaction term implies that for longer relations, the measured effect of collateral is positive. This is again consistent with the net empirical dominance of the ex post theories for seasoned customers.

The independent effect of relation length is essentially zero in Column 1 when no interaction term is included. The negative coefficient for relation length in Columns 2 and 3 implies that, when no collateral is pledged, firms with longer relations are less likely to have nonperformance problems, consistent with expectations that such borrowers are less risky.[20] With respect to the other control variables, we find that partnerships and corporations are more likely to have loan performance problems than sole proprietorships. This is consistent with limited liability playing a role in default decisions. Regarding contract terms, it appears that installment loans and loans with longer maturities are associated with a higher incidence of repayment problems. Larger loans, by contrast, are negatively associated with ex post nonperformance.

As a robustness check, we reestimate the specifications reported in Table 5 separately for loans with contract amounts above and below $75,000. Results are similar to those reported in Table 5. We also reestimate the specifications reported in Table 5 after dropping all loans that appear to be renegotiations of previous loans. (Loans designated in the registry as renegotiations have already been excluded from this analysis.) Including such loans could bias the estimated relations between ex post nonperformance and collateral. This bias would arise in situations in which the borrower became distressed and the bank demanded that collateral is pledged but recorded the adjustment as a new loan (see, e.g., discussions in Smith, 1993; DeAngelo, DeAngelo, and Wruck, 2002). To identify such situations, we look for loans that are originated right after another loan at the same bank terminates. We identify 5,962 such loans, of which only 424 have collateral added. Reestimating Eq. (3) without these loans has no effect on our results.

6. Conclusions

The theoretical literature offers two broad classes of theories about why borrowers pledge collateral. The first set of theories motivates collateral as a way for good borrowers to signal their quality under conditions of ex ante private information. The second set of theories explains collateral as an optimal response to ex post contract frictions such as moral hazard. A growing body of literature that empirically tests these models and the recent financial crisis have raised significant academic and

[20] Although borrowers with longer relations appear to be less risky on average, they do not pay lower risk premiums (see Columns 2 and 3 in Table 4). This is consistent with the extraction of information rents as in Sharpe (1990) and von Thadden (2004). See also Ioannidou and Ongena (2010) for evidence consistent with this hypothesis.

Table 5
Determinants of ex post nonperformance.
This table reports probit regressions for *Ex_Post_Nonperformance*, a dummy variable that equals one if a loan is 30+ days overdue anytime after its origination or if it is downgraded to the default status (i.e., given a rating of 5). Under the heading "Probit coefficients," we report the estimated coefficients of the three probit specifications. Standard errors, corrected for heteroskedasticity, are reported between brackets. Under the heading "Marginal effects," we report the change in probability of pledging collateral for each one of the independent variables. For continuous variables we report the effect for an infinitesimal change in each independent variable and for dummy variables we report the estimated effect of a change from zero to one. P0 is the predicted probability of ex post nonperformance, evaluated at the mean of all independent variables. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variables	Probit coefficients			Marginal effects		
	(1)	(2)	(3)	(1)	(2)	(3)
Loan characteristics						
Collateral	0.362***	−0.044	−0.069	0.041***	−0.004	−0.006
	[0.031]	[0.080]	[0.082]	[0.004]	[0.007]	[0.007]
Collateral∗Rel_Length		0.153***	0.154***		0.015***	0.015***
		[0.028]	[0.028]		[0.003]	[0.003]
Relation characteristic						
Rel_Length	0.01	−0.051***	−0.041***	0.001	−0.005***	−0.004**
	[0.016]	[0.019]	[0.019]	[0.002]	[0.002]	[0.002]
Firm characteristics						
Partnerships	0.249***	0.25***	0.263***	0.028***	0.028***	0.029***
	[0.052]	[0.052]	[0.053]	[0.007]	[0.007]	[0.007]
Corporations	0.127***	0.126***	0.16***	0.012***	0.012***	0.014***
	[0.043]	[0.043]	[0.044]	[0.004]	[0.004]	[0.004]
Other	−0.086	−0.075	−0.009	−0.008	−0.007	−0.001
	[0.107]	[0.107]	[0.107]	[0.009]	[0.009]	[0.010]
Other loan characteristics						
Installment Loan			0.188***			0.018***
			[0.035]			[0.003]
Loan Amount			−0.054***			−0.005***
			[0.009]			[0.001]
Maturity			0.076***			0.007***
			[0.021]			[0.002]
Industry, *Region*, *Bank*, and *Time* dummy variables	Included	Included	Included	Included	Included	Included
P0				0.047	0.046	0.045
Pseudo R-squared	0.11	0.11	0.12	0.11	0.11	0.12
Number of observations	28,729	28,729	28,729	28,729	28,729	28,729

policy interest in understanding the role of collateral in debt contracts.

This paper improves upon the empirical literature by using data from the Bolivian public credit registry that provides important risk information about the borrower that is not known to the lender. Thus, we have both private and public information about the firm. Using this information structure, we are able to construct measures of both observed and unobserved risk and, hence, more effectively test the two sets of collateral theories. The data also allow us to explore the role of banking relations and how information gleaned from relations reduces private information.

We present results that suggest roles for both sets of theories, although the ex ante private information theories appear to hold only for customers with relatively short relations that are relatively unknown to the lender. The data also suggest that the ex post theories tend to empirically dominate for firms with long relations, where private information is less important.

Our analysis represents an important contribution to the literature seeking to understand the motivation for collateral in debt contracts. First, the issue has clearly been on the minds of market participants and policy makers in places such as Japan and the United States owing to significant shocks to collateral values. Second, we use credit registry data that allow us to produce clean measures of private and public information, as well as providing a rich set of controls at the loan and bank level and bank and time fixed effects to account for unobserved bank heterogeneity and changes in the lending environment, respectively. Our approach might also be relevant to World Bank efforts to encourage the establishment of the development of credit registries in the developing world. Our findings suggest that the information provided by such registries might be useful in eliminating the need for costly collateral.

Appendix A

Our empirical analysis rests on the maintained assumption that past loan performance is predictive of future loan performance. That is, borrowers with past repayment problems are more likely to become delinquent or default on future loans. To investigate whether this assumption holds for our sample, we examine how our four measures of risk based on past repayment histories relate to the performance of new loans using the probit model

$$P(Ex\_Post\_Nonperformance_{ijt}) = h(Observed\_Risk_{ijt},\ Unobserved\_Risk_{ijt},\ Rel\_Length_{ijt},\ Firm_{ijt},\ Loan_{ij},\ \alpha_j,\ \gamma_t), \quad (4)$$

Table A1
Past performance is predictive of future performance.

This table reports probit regressions for *Ex_Post_Nonperformance*, a dummy variable that equals one if a loan is 30+ days overdue anytime after its origination or if it is downgraded to the default status (i.e., given a rating of 5). Under the heading "Probit coefficients," we report the estimated coefficients of the two probit specifications. Standard errors, corrected for heteroskedasticity, are reported between brackets. Under the heading "Marginal effects," we report the change in probability of pledging collateral for each one of the independent variables. For continuous variables we report the effect for an infinitesimal change in each independent variable and for dummy variables we report the estimated effect of a change from zero to one. P0 is the predicted probability of ex post nonperformance, evaluated at the mean of all independent variables. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variables	Probit coefficients		Marginal effects	
	(1)	(2)	(1)	(2)
Observed risk				
Default_Observable_Registry	0.081	0.05	0.008	0.004
	[0.189]	[0.190]	[0.019]	[0.018]
Npl_Observable_Registry	0.512***	0.546***	0.067***	0.071***
	[0.045]	[0.046]	[0.008]	[0.008]
Npl_Observable_Relation	0.729***	0.728***	0.109***	0.105***
	[0.038]	[0.038]	[0.008]	[0.008]
Unobserved risk				
Npl_Unobservable	0.192***	0.216***	0.019***	0.021***
	[0.035]	[0.036]	[0.004]	[0.004]
Relation characteristic				
Rel_Length	−0.073***	−0.06***	−0.007***	−0.005***
	[0.015]	[0.016]	[0.001]	[0.001]
Borrower characteristics				
Partnerships	0.169***	0.178***	0.017***	0.017***
	[0.054]	[0.055]	[0.006]	[0.006]
Corporations	0.09**	0.125***	0.008**	0.01***
	[0.044]	[0.045]	[0.004]	[0.004]
Other	−0.091	−0.02	−0.008	−0.002
	[0.110]	[0.110]	[0.008]	[0.009]
Loan characteristics				
Installment Loan		0.195***		0.017***
		[0.036]		[0.003]
Loan Amount		−0.055***		0.005***
		[0.009]		[0.001]
Maturity		0.12***		0.01***
		[0.021]		[0.002]
Industry, *Region*, *Bank*, and *Time* dummy variables	Included	Included	Included	Included
P0			0.042	0.04
Pseudo R-squared	0.152	0.166	0.152	0.166
Number of observations	28,729	28,729	28,729	28,729

where $Ex\_Post\_Nonperformance_{ijt}$ is a dummy variable that equals one if the loan is 30+ days overdue anytime after origination or if it enters default status. All other variables are defined as in Eq. (1). The model is estimated using the 28,758 loans that were originated prior to the last six months of the sample and matured before the end of the sample.

The estimation results are presented in Table A1.[21] Column 1 reports results from a benchmark specification without loan characteristics. Column 2 reports results from a model including loan characteristics. Under the heading "Probit coefficients," we report the estimated coefficients of the two probit specifications. Under the heading "Marginal effects," we report the change in probability of ex post nonperformance for each one of the independent variables, holding all other independent variables at their sample means. For continuous variables, we report the effect for an infinitesimal change in the variable. For dummy variables, we report the estimated effect of a change from zero to one.

In both specifications, all four indicators of borrower risk based on past repayment histories are positively correlated with repayment problems on the new loan. This suggests that past performance is predictive of future performance. Moreover, considering that the predicted probability of the new loan becoming nonperforming (at the mean of all dependent variables, P0) is around 4%, the estimated marginal effects of these four risk indicators are large.

[21] Results are similar if we estimate the model separately for loans above or below $75,000.

References

Aghion, P., Bolton, P., 1997. A theory of trickle-down growth and development. Review of Economic Studies 64, 151–172.
Albuquerque, R., Hopenhayn, H.A., 2004. Optimal lending contracts and firm dynamics. Review of Economic Studies 71, 285–315.
Banerjee, A.V., Newman, A.F., 1993. Occupational choice and the process of development. Journal of Political Economy 101, 274–298.
Berger, A.N., Espinosa-Vega, M.A., Frame, W.S., Miller, N.H. Why do borrowers pledge collateral? New empirical evidence on the role of

asymmetric information, Journal of Financial Intermediation, in press, doi:10.1016/j.jfi.2010.01.001.
Berger, A.N., Udell, G.F., 1990. Collateral, loan quality, and bank risk. Journal of Monetary Economics 25, 21–42.
Berger, A.N., Udell, G.F., 1995. Relationship lending and lines of credit in small firm finance. Journal of Business 68, 351–381.
Bernanke, B.S., Gertler, M., 1989. Agency costs, net worth, and business fluctuations. American Economic Review 79, 14–31.
Bernanke, B.S., Gertler, M., 1990. Financial fragility and economic performance. Quarterly Journal of Economics 105, 87–114.
Besanko, D., Thakor, A.V., 1987a. Collateral and rationing: sorting equilibria in monopolistic and competitive credit markets. International Economic Review 28, 671–689.
Besanko, D., Thakor, A.V., 1987b. Competitive equilibrium in the credit market under asymmetric information. Journal of Economic Theory 42, 167–183.
Bester, H., 1985. Screening vs. rationing in credit markets with imperfect information. American Economic Review 75, 850–855.
Bester, H., 1987. The role of collateral in credit markets with imperfect information. European Economic Review 106, 60–75.
Blackwell, D.W., Winters, D.B., 1997. Banking relationships and the effect of monitoring on loan pricing. Journal of Financial Research 20, 275–289.
Boot, A.W.A., Thakor, A.V., 1994. Moral hazard and secured lending in an infinitely repeated credit market game. International Economic Review 35, 899–920.
Boot, A.W.A., Thakor, A.V., Udell, G.F., 1991. Secured lending and default risk: equilibrium analysis, policy implications and empirical results. Economic Journal 101, 458–472.
Boyd, J.H., Smith, B.D., 1994. The equilibrium allocation of investment capital in the presence of adverse selection and costly state verification. Economic Theory 3, 427–451.
Brick, I.E., Palia, D., 2007. Evidence of jointness in the terms of relationship lending. Journal of Financial Intermediation 16, 452–476.
Campion, A., 2001. Client information sharing in Bolivia. Journal of Microfinance 3, 45–63.
Chakraborty, A., Hu, C.X., 2006. Lending relationships in line-of-credit and nonline-of-credit loans: evidence from collateral use in small business. Journal of Financial Intermediation 15, 86–107.
Chan, Y.-S., Thakor, A.V., 1987. Collateral and competitive equilibria with moral hazard and private information. Journal of Finance 42, 345–363.
Cole, R.A., 1998. The importance of relationships to the availability of credit. Journal of Banking and Finance 22, 959–977.
Cooley, T., Marimon, R., Quadrini, V., 2004. Aggregate consequences of limited contract enforceability. Journal of Political Economy 112, 817–847.
de Janvry, A., Sadoulet, E., McIntosh, C., Wydick, B., Luoto, J., Gordillo, G., Schuetz, G., 2003. Credit bureaus and the rural microfinance sector: Peru, Guatemala, and Bolivia. Joint Project with the University of California and the Food and Agricultural Office for Latin America, Berkeley, CA, and Santiago, Chile. Available at: ⟨http://are.berkeley.edu/~sadoulet/papers/Final_report_FAO.pdf⟩.
DeAngelo, H., DeAngelo, L., Wruck, K.H., 2002. Asset liquidity, debt covenants, and managerial discretion in financial distress: the collapse of L.A. Gear. Journal of Financial Economics 64, 3–34.
Degryse, H., Van Cayseele, P., 2000. Relationship lending within a bank-based system: evidence from European small business data. Journal of Financial Intermediation 9, 90–109.
Djankov, S., McLiesh, C., Shleifer, A., 2007. Private credit in 129 countries. Journal of Financial Economics 84, 299–329.
Elsas, R., Krahnen, J.P., 2000. Collateral, default risk, and relationship lending: an empirical study on financial contracting. Discussion Paper 2540, Centre for Economic Policy Research, London, England.
Gale, D., Hellwig, M., 1985. Incentive-compatible debt contracts: the one-period problem. Review of Economic Studies 52, 647–663.
Gan, J., 2007. Collateral, debt capacity, and corporate investment: evidence from a natural experiment. Journal of Financial Economics 85, 709–734.
Gonas, J.S., Highfield, M., Mullineaux, D., 2004. When are commercial loans secured? Financial Review 39, 79–99.
Greenbaum, S.I., Kanatas, G., Venezia, I., 1989. Equilibrium loan pricing under the bank–client relationship. Journal of Banking and Finance 13, 221–235.
Holmstrom, B., Tirole, J., 1997. Financial intermediation, loanable funds, and the real sector. Quarterly Journal of Economics 62, 663–691.
Harhoff, D., Korting, T., 1998. Lending relationships in Germany: empirical evidence from survey data. Journal of Banking and Finance 22, 1317–1353.
Ioannidou, V., Ongena, S., 2010. Time for a change: loan conditions and bank behavior when firms switch banks. Journal of Finance 65, 1847–1877.
Jimenez, G., Salas, V., Saurina, J., 2006. Determinants of collateral. Journal of Financial Economics 81, 255–281.
Jimenez, G., Saurina, J., 2004. Collateral, type of lender and relationship banking as determinants of credit risk. Journal of Banking and Finance 28, 2191–2212.
John, K., Lynch, A.W., Puri, M., 2003. Credit ratings, collateral, and loan characteristics: implications for yield. Journal of Business 76, 371–409.
Kiyatoki, N., Moore, J., 1997. Credit cycles. Journal of Political Economy 105, 211–248.
Leeth, J.D., Scott, J., 1989. The incidence of secured debt: evidence from the small business community. Journal of Financial and Quantitative Analysis 24, 379–394.
Lehmann, E., Neuberger, D., 2001. Do lending relationships matter? Evidence from bank survey data in Germany. Journal of Economic Behavior and Organization 45, 339–359.
Liberti, J., 2009. Uncovering collateral constraints. Unpublished Working Paper, DePaul University, Chicago, IL.
Machauer, A., Weber, M., 1998. Bank behavior based on internal credit ratings of borrowers. Journal of Banking and Finance 22, 1355–1383.
Menkhoff, L., Neuberger, D., Suwanaporn, C., 2006. Collateral-based lending in emerging markets: evidence from Thailand. Journal of Banking and Finance 30, 1–21.
Ono, A., Uesugi, I., 2009. Role of collateral and personal guarantees in relationship lending: evidence from Japan's SME loan market. Journal of Money, Credit, and Banking 41, 935–960.
Petersen, M.A., Rajan, R.G., 1994. The benefits of lending relationships: evidence from small business data. Journal of Finance 110, 407–443.
Rajan, R.G., 1992. Insiders and outsiders: the choice between informed and arm's-length debt. Journal of Finance 47, 1367–1400.
Sharpe, S.A., 1990. Asymmetric information, bank lending, and implicit contracts: a stylized model of customer relationships. Journal of Finance 45, 1069–1087.
Smith Jr., C.W., 1993. A perspective on accounting-based debt covenant violations. Accounting Review 68, 289–303.
Smith Jr., C.W., Warner, J.B., 1979a. Bankruptcy, secured debt, and optimal capital structure: comment. Journal of Finance 34, 247–251.
Smith Jr., C.W., Warner, J.B., 1979b. On financial contracting: an analysis of bond covenants. Journal of Financial Economics 7, 117–161.
Sirtaine, S., Skamnelos, I., Frank, S., 2004. Bolivia: Challenges in the Corporate and Banking Sectors. World Bank, Washington, DC.
Stiglitz, J.E., Weiss, A., 1981. Credit rationing in markets with imperfect information. American Economic Review 71, 393–410.
Townsend, R.M., 1979. Optimal contracts and competitive markets with costly state verification. Journal of Economic Theory 21, 265–293.
von Thadden, E.L., 2004. Asymmetric information, bank lending, and implicit contracts: the winner's curse. Finance Research Letters 1, 11–23.
Williamson, S.D., 1986. Costly monitoring, financial intermediation and equilibrium credit rationing. Journal of Monetary Economics 18, 159–179.

Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Financial Economics 76 (2005) 237–269
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/econbase

Does function follow organizational form? Evidence from the lending practices of large and small banks ☆

Allen N. Berger[a], Nathan H. Miller[b], Mitchell A. Petersen[c], Raghuram G. Rajan[d], Jeremy C. Stein[e,*]

[a]*Board of Governors of the Federal Reserve System and Wharton Financial Institutions Center*
[b]*University of California, Berkeley, USA*
[c]*Northwestern University and NBER, USA*
[d]*University of Chicago and NBER, USA*
[e]*Department of Economics, Harvard University, Cambridge, MA 02138 and NBER, USA*

Received 24 April 2003; received in revised form 31 March 2004; accepted 11 June 2004
Available online 5 February 2005

Abstract

Theories based on incomplete contracting suggest that small organizations have a comparative advantage in activities that make extensive use of "soft" information. We provide evidence consistent with small banks being better able to collect and act on soft information than large banks. In particular, large banks are less willing to lend to informationally "difficult" credits, such as firms with no financial records. Moreover, after controlling for the endogeneity of bank-firm matching, we find that large banks lend at a greater distance, interact more

☆The opinions in this paper do not necessarily reflect those of the Federal Reserve Board or its staff. This work has been supported by the National Science Foundation (Rajan, Stein), and the George J. Stigler Center for Study of the State and Economy (Rajan). We are grateful to Phil Ostromogolsky for outstanding research assistance. Thanks also to seminar participants at Yale, the Federal Reserve Banks of New York and Atlanta, Tulane, Babson, Dartmouth, Washington University, the University of Illinois, the Federal Reserve Bank of Chicago Bank Structure Conference, the NBER, and the Western Finance Association meetings, as well as to Abhijit Banerjee, Michael Kremer, David Scharfstein, Andrei Shleifer, David Smith, Greg Udell, Christopher Udry, James Weston, Michael Whinston, and the referees for helpful comments and suggestions.

*Corresponding author.
E-mail address: stein2@fas.harvard.edu (J.C. Stein).

doi:10.1016/j.jfineco.2004.06.003

impersonally with their borrowers, have shorter and less exclusive relationships, and do not alleviate credit constraints as effectively.

JEL classification: G21; L22

Keywords: Organizations; Soft information; Banking

1. Introduction

One of the most enduring questions in economics was posed by Coase (1937): What determines the boundaries of the firm? The question is perhaps most often framed in terms of *vertical* integration—i.e., when can it make sense for upstream and downstream activities to be combined under the roof of a single firm? But one can also ask about the circumstances under which *horizontal* integration creates value. A good present-day illustration of this version of the question comes from the commercial banking industry, where ongoing consolidation raises the issue of whether the resulting large banks will behave differently than the small banks that they are displacing.

A partial answer to Coase's question comes from the work on transaction-cost economics of Williamson (1975, 1985) and Klein et al. (1978). These authors focus on the hold-up problems that can accompany market transactions, and argue that such problems can be mitigated by having the firm, rather than the market, mediate trade. While this approach is helpful in identifying the advantages of integration (i.e., a reduction in market hold-up problems), it is less clear on the disadvantages. As such, it is somewhat of a one-sided theory—unless one invokes factors outside the model, like unspecified "costs of bureaucracy," it has the awkward implication that efficiency would be best served by placing all of the economy's assets inside a single firm.

The disadvantages of integration emerge much more clearly in the framework of Grossman and Hart (1986), Hart and Moore (1990), and Hart (1995), henceforth GHM. At its most general level, the central insight of the GHM paradigm is that, in a world of incomplete contracts, agents' ex ante incentives are shaped by the extent to which they have control or authority over physical assets. Thus, for example, if firm A acquires firm B, the manager who was previously CEO of firm B might reduce his level of ex ante effort now that he is subordinate to the CEO of firm A and no longer has full control rights over B's assets; herein lies the potential cost of integration.

The GHM paradigm has strongly influenced subsequent work on the theory of the firm. But it has proved challenging to construct sharp empirical tests of the theory. As discussed in Whinston (2001), this is in part due to the fact that the predictions can be sensitive to specific assumptions, such as the nature of the non-contractible investments that need to be made ex ante. Another difficulty is that GHM focus on ownership of physical assets as the exclusive source of power and incentives in the firm, thereby abstracting from other considerations that might be present in a richer, more empirically realistic model. These considerations include differentially informed agents as in Aghion and Tirole (1997), incentive structures as in Holmstrom and Milgrom (1994) and Holmstrom (1999), or access to critical resources as in Rajan and Zingales (1998, 2001).

One strategy for dealing with these problems is to not take the original GHM models too literally as a basis for empirical testing, and to work instead with "second-generation" models that build on the basic GHM insights but that are more tailored to delivering clear-cut comparative static predictions, either for a specific type of investment or in a particular institutional setting. This strategy is followed by Baker and Hubbard (2003), whose work centers on the trucking industry and the question of whether drivers should own the trucks they operate, as well as by Simester and Wernerfelt (2003), who look at the ownership of tools in the carpentry industry.

In this paper, we take a broadly similar approach. Unlike the above-mentioned authors, however, our focus is not on how differences in technology influence the ownership of assets, but rather on how the nature of an organization affects both the way it does business and the kinds of activities that it can efficiently undertake. Specifically, we attempt to understand whether small organizations are better at carrying out certain tasks than large organizations. In this regard, our work is closer to Mullainathan and Scharfstein (2001), who document how producers of a particular chemical that are integrated with the downstream users of the chemical have investment behavior that differs from that of stand-alone producers.

Our starting point is the model in Stein (2002). This model adopts the basic GHM insight that the allocation of control affects incentives, but it does so in a setting that is more specific, and thus yields sharper empirical predictions. The predictions have to do with the differing incentives that are created in large and small firms for the production and use of various kinds of information. The model implies that small firms are at a comparative advantage in evaluating investment projects when the information about these projects is naturally "soft" and cannot be credibly communicated from one agent in the firm to another. In contrast, large firms do relatively well when information about investment projects can be easily "hardened" and passed along within the hierarchy.

This model applies naturally to the banking industry, where information is critical to the activity of lending. The model suggests that large banks will tend to shy away from small-business lending, because this is an activity that relies especially heavily on the production of soft information, something that is not their strong suit. For example, consider a loan officer trying to decide whether or not to extend credit to a small start-up company that does not have audited accounting statements. The best the loan officer may be able to do is to spend time with the company president in an effort to determine whether she is honest, prudent, and hardworking—i.e., the classic candidate for a "character loan." However, given that this information is soft and cannot be verifiably documented in a report that the loan officer can pass on to his superiors, the model predicts (as is explained in more detail below) that his incentives to produce high-quality information are weak when he works inside a large bank.

By contrast, when dealing with a larger company that has a well-documented track record, the decision to extend credit can be based more heavily on verifiable information, such as the company's income statements, balance sheet, and credit rating. In this case, the model suggests that a large bank will have no problem—indeed, it may do better—at providing incentives for information production.

To test this theory, we make use of a data set on small-business lending that has information not only about the small firms in the sample, but also about their primary bank lenders and the nature of the relationship between the two. The data thus allow us to investigate a number of hypotheses about how the "technology" of lending depends on variables such as bank size. If, as the theory suggests, large banks are at a comparative disadvantage in the production and use of soft information, one would expect this to influence their methods of lending.

We develop six basic pieces of evidence to support this case. First, and most simply, we find that bigger banks are more apt to lend to firms that are larger or that have better accounting records (a good example of hard information). Second, controlling for firm and market characteristics, we find that the physical distance between a firm and the branch office that it deals with increases with the size of the bank. This is consistent with the notion that large banks rely less on the sort of soft information that is typically available through personal contact and observation. Third and relatedly, we find that firms do business with large banks in more impersonal ways—i.e., they meet less often with their banker and instead communicate more by mail or phone.

Of course, a firm chooses the bank from which it borrows. That is, the match between a firm and its bank is to some extent endogenous, and is likely to be related to firm characteristics. Indeed, our first finding—that bigger banks match up with firms with better accounting records—is evidence of just this endogeneity. This suggests that we need to proceed carefully if, as in our second and third findings, we want to use bank size as a right-hand-side variable to explain certain aspects of the lending relationship. For example, perhaps large banks deal with their customers more impersonally not because they are incompetent at personal interaction, but because they tend to match with a type of customer for whom personal interaction is less appropriate.

To deal with this potential endogeneity problem, we try instrumenting for bank size with two variables: (i) the median size of *all* banks (weighted by number of branches) in the market where the firm is located, and (ii) a regulatory variable that measures how permissive the firm's state has been with respect to branching. Intuitively, if a firm borrows from a large bank because it is located in a market where there are only large banks (say because regulation has not artificially constrained bank size), this does not reflect an endogenous choice on the part of the firm, but rather an exogenous, geographically imposed limitation. We find that when we take this instrumental-variables (IV) approach, the estimated effect of bank size on distance and on the extent of impersonal communication is even larger than when we do not correct for endogeneity.

Our fourth and fifth findings are that bank-firm relationships tend to be stronger—both more long-lived and more exclusive—when the firm in question borrows from a small bank. These findings also emerge both with and without using IV, but again are more pronounced when an IV approach is employed. They are exactly what one would expect based on the theory, given that the soft information produced by small banks is more likely than hard information to be non-transferable. In other words, the theory suggests that small-bank lending should fit more closely with the kind of model in Rajan (1992), where accumulated soft information binds a borrower to its bank over time.

The sixth and final part of our empirical work is to test whether bank size affects the availability of credit to small businesses. If small firms need lenders that are willing to go deeper and acquire soft information, then we would expect those that are forced to go to large banks to be particularly credit constrained. One measure of the degree to which a firm is rationed by financial institutions is the amount of expensive trade credit it relies on (Petersen and Rajan (1994) and Fisman and Love (2003)). We find that all else equal, a firm that borrows from a larger bank is more prone to repay its trade credit late.

Interestingly, this last result holds *only* when we instrument for bank size. When firms are forced to borrow from large banks because there are no small banks around, they seem to face credit constraints—this is what the IV version of the regression tells us. At the same time, an ordinary-least-squares regression of credit constraints on bank size reveals an offsetting effect due to the endogeneity bias: those firms that are by nature the most difficult credits tend to match with smaller banks, as the theory would suggest.

While our empirical tests are primarily motivated by a model in the control-rights genre, it is important to note that some of the same predictions about the effects of bank size follow from other types of agency models. To take a leading example, Brickley et al. (2003) observe that officers and directors in their sample of small Texas banks own an average of nearly 70% of the stock of these banks. They then go on to propose a theory of small-bank/big-bank differences based on explicit incentive-contracting considerations. In particular, they argue that since managers of small banks have higher-powered ownership incentives, they will devote more effort to soft-information collection, and can be trusted to use this information in a way that is consistent with shareholder objectives. This differs from Stein's (2002) theory, which emphasizes the incentive effects of control rights rather than of direct share ownership.

Our view is that these two types of theories are broadly complementary, and it would be a mistake to try to argue that our basic empirical findings are solely the product of one mechanism or the other. Nevertheless, the two theories have some divergent implications, which in principle allow for a degree of separation.

Section 2 reviews both theories and fleshes out our main hypotheses more fully. Section 3 introduces our data set. Section 4 describes our empirical results. Section 5 discusses how our work fits with the related banking literature, and Section 6 concludes.

2. Hypothesis development

2.1. Overview of the theory

The logic of Stein's (2002) model can be sketched with an example. Imagine a loan officer in Little Rock who is responsible for deciding which small-business loans are worth making. The quality of the loan officer's judgment will depend on how good a job he has done in producing soft information, which in turn will be a function of his incentives. In the limiting case of a very small bank, the loan officer is also president of the bank, and has the authority to allocate the bank's funds as he sees fit. Given that he can count on having some capital to work with, he knows that his research efforts will not be wasted, and hence his incentives to do research are relatively strong. In other words, the decentralization inherent in a small bank rewards an agent who develops expertise by ensuring the availability of capital with which to lever that expertise.

In contrast, if the Little Rock loan officer is part of a large multi-branch hierarchy, the following problem arises. Suppose that he spends a lot of effort learning about local prospects. But then somebody higher up in the organization decides that overall lending opportunities are better in Tulsa, and cuts the capital allocation for Little Rock. In this case, because he doesn't get a chance to act on the soft information that he has produced, and because he cannot credibly pass it on, the loan officer's research effort goes to waste. Ex ante, this implies that the loan officer does less research in a hierarchical setting. Here, the authority to allocate capital is separated from expertise—i.e., the Little Rock loan officer can be left with no capital to work with—which dilutes the incentives to become an expert. This can be thought of as a specific manifestation of the key GHM idea that taking control away from an agent tends to weaken the agent's incentives.[1]

To further bring out the intuition of the model with soft information, consider this question: All else equal, will a large banking organization be better at making small-business loans if it is set up as single legal entity, or as a multi-bank holding company with a number of legally distinct subsidiaries? Several authors (e.g., Keeton, 1995; DeYoung et al., 1998) hypothesize that the multi-bank holding company structure is particularly inimical to small-business lending, because it adds extra layers of bureaucracy. However, Stein (2002) argues that just the opposite may be the case. To the extent that this structure makes it harder to move capital across the different subsidiaries, it can act as a partial precommitment by the CEO to run a decentralized operation—i.e., to not reduce individual agents' capital allocations. This should improve their incentives to gather soft information, and thereby benefit small-business lending.

[1]Aghion and Tirole (1997) also argue that agents' incentives can be blunted when they are in a hierarchy. A critical distinction is that in Stein (2002), a hierarchical structure *need not* weaken incentives—indeed, it only does so when information is soft. Thus the models have quite different empirical implications: the Aghion–Tirole model does not say anything about why large banks might be at more of a disadvantage with small-business loans than with credit cards or mortgages.

The model works very differently when the information produced by agents can be hardened and passed on to their superiors, as might be the case with the output from a credit-scoring model. Now, large banks might actually generate more investigative effort than small banks. This is because with hard information, agents can become advocates for their units—a Little Rock loan officer working inside a large bank who produces verifiable evidence showing that lending opportunities in his area are strong can increase his capital allocation. Here, separating authority from expertise actually improves research incentives, as lower-level managers struggle to produce enough information to convince their superiors that they should get a larger share of the bank's overall capital budget. Similarly, Rajan and Zingales (1998) observe that withholding ownership can in some cases spur effort by encouraging competition for power.

Although the explicit distinction between soft and hard information that Stein emphasizes is not typically drawn in the applied banking literature, it corresponds closely to the oft-discussed dichotomy between "relationship" lending and "transactions-based" lending (see, e.g., Berger and Udell, 2002). Moreover, it is a common hypothesis in this line of work that large banks will be at a disadvantage when it comes to relationship lending, but will do better with respect to transactions-based lending. For example, Berger et al. (1999) argue that "because of….organizational diseconomies…large complex financial institutions….would reduce services …to those customers who rely on relationships" (pp. 165–166).

2.2. Testable implications

2.2.1. The choice of bank

The most basic implication of the theory is that small banks have a comparative advantage in making loans based on soft information, while large banks have a comparative advantage in making loans based on hard information. This suggests that a firm about which there is more hard information should tend to borrow from a larger bank. One potential proxy for whether there is hard information about a firm is its size, since, e.g., large firms are more able to afford the fixed costs associated with regular audits. Of course, there can be other reasons why large firms and large banks go together. However, our data also allow us to infer whether a given firm keeps accounting records. This could serve as an alternative proxy for hard information, and we would therefore predict that firms with accounting records are more likely to borrow from larger banks.

2.2.2. The endogeneity of bank size and our instrumenting strategy

All the hypotheses that follow relate bank size to various aspects of the bank-firm lending relationship. In other words, we want to use bank size as a right-hand-side variable to explain the nature of the lending technology. But since firms can to a degree choose their banks—as we have just emphasized—there is an obvious endogeneity problem. In particular, a firm characteristic for which we have not controlled could explain why the firm chooses a bank of a certain size, as well as the aspect of the relationship we are interested in. For example, an entrepreneur with an

MBA might be better able to get a hearing from a loan officer in a large bank. This entrepreneur might also find it easier to generate spreadsheet reports that reduce the need for personal visits to the bank. Thus, he might be more apt to borrow at a distance, and to communicate with the bank impersonally. If so, we would see large banks lending impersonally and at a distance, but this would not necessarily reflect a causal consequence of bank size.

To address this potential bias, we need one or more instruments that are correlated with a firm's propensity to be matched with a bank of a particular size, but that are uncorrelated with characteristics of the firm that might influence the nature of the lending relationship. In our baseline specifications, we use two instruments: (i) the log of the median size of *all* the banks in the Metropolitan Statistical Area (MSA) or rural county in which the firm is located (weighted by the number of branches), and (ii) the fraction of the previous ten years during which the firm's state was neither a unit banking nor limited branching state. The idea is that if a firm is located in a state where regulation has not constrained bank size, and hence where large banks dominate its market, the firm will be pushed—independent of its own characteristics—in the direction of a large bank. We can then examine how this forced match shapes the bank-firm relationship.

Although our median-bank-size instrument varies at the level of the city or rural county, and our regulatory instrument varies only at the state level, the two are closely linked, with a univariate correlation of 0.472. Not surprisingly, states that have been permissive with respect to branching tend to have larger banks across all of their individual markets. In spite of this commonality, however, one might argue that the state-level regulatory variable is a purer instrument. Perhaps within a given state, some markets have certain attributes that tend to attract both banks of a certain size and firms with particular characteristics. For example, a vibrant big-city economy might draw both large banks and MBA-trained entrepreneurs.

An alternative estimation strategy that helps to address this critique is to dispense with the median-bank-size variable and to use the regulatory variable as the *only* instrument for bank size. This approach, which we experiment with below, is more conservative, but also considerably less powerful, because it makes use only of across-state variation, and loses the within-state across-market variation. Nevertheless, it leads to point estimates that are remarkably similar to those from our baseline instrumenting technique, although the standard errors are of course somewhat higher.

2.2.3. The effect of bank size on distance and mode of interaction

Coval and Moskowitz (2001) demonstrate the importance of physical distance for information-gathering, documenting that money managers do better when investing in the stocks of nearby companies. Similarly, being close to one's customers is likely to facilitate a loan officer's collection of soft information, but to have little impact on his ability to gather hard information. What we have in mind is that one important way for the loan officer to gather soft information is through face-to-face interaction with a potential borrower. Being nearby might also help the loan officer to better understand the nuances of the local business environment. Hard information, on the other hand, can by definition be easily summarized in a report, and hence can be faxed or emailed anywhere, making distance essentially irrelevant.

Now think of a firm that wants to borrow. If it is forced to choose among large banks (because, say, no small banks are around), we would not expect the firm to limit itself to those that are close, because large banks are unlikely to invest in acquiring soft information, making the lending technology more distance-independent. We would also expect the mode of communication between the firm and the bank to be more impersonal. By contrast, if only small banks are around and the firm is informationally opaque, we would expect it to pick a nearby bank, because small banks' information acquisition is sensitive to the "shoe-leather" cost of personal visits. We would also expect the contact between the firm and bank to be more personal in nature.

2.2.4. The effect of bank size on relationship length and exclusivity

If our findings about distance and mode of interaction reflect the fact that small banks are better at using soft information, we should see this manifested in two further ways. First, small banks should sustain longer relationships with their borrowers. The soft information that a small bank has gathered over time should give it a comparative advantage in providing its client firm with good lending terms. Moreover, because this soft information is not easily transferable by the firm, the banker might have a certain degree of market power (see Sharpe, 1990; Rajan, 1992), which would further tie the firm to the bank. If, on the other hand, a firm's relationship with a large bank is based on hard information, which is easily communicated to potential new lenders, the additional benefits of staying with the same lender, or the switching costs of moving to a new one, are likely to be lower. The length of time that a firm and its bank have dealt with each other should thus decrease with bank size.

A second implication, which follows from similar reasoning, is that the likelihood that a relationship between a firm and its bank is an exclusive one—i.e., that the bank is the firm's only lender—should also decrease with bank size. In other words, their greater reliance on soft information suggests that smaller banks should form both longer and more exclusive relationships with their customers.

2.2.5. The effect of bank size on credit availability

Since we argue that small banks form stronger, more information-intensive bonds with their borrowers, we might also expect them to do a better job of easing these firms' credit constraints. If we can document evidence consistent with this prediction, we will have identified an important "real" effect of bank size that would seem to be particularly difficult to explain away with alternative theories.

To form an operational measure of credit constraints, we follow Petersen and Rajan (1994) and look at the fraction of a firm's trade credit that is paid late. As Petersen and Rajan argue, stretching trade credit is a very expensive way to obtain finance, and a firm is likely to do so only when rationed by institutional lenders. So the final prediction of our theory is that firms should repay a higher fraction of their trade credit late if they borrow from larger banks. This is perhaps the test for which it is most critical to correct for the endogeneity of the firm's choice of bank, as one would expect particularly difficult credits to choose small banks. Without instrumenting for bank size, the test would therefore be biased against finding that small banks improve credit availability.

2.3. Alternative theories

As noted in the introduction, many of the above predictions about small-bank/big-bank differences can also be motivated in the context of a model based on explicit incentive-contracting considerations. Brickley et al. (2003) point out that managers of small banks tend to have higher-powered ownership incentives, and there are a variety of reasons to expect that such well-aligned incentives might facilitate a reliance on soft information. In spite of the similarities, however, one can imagine tests that might discriminate between the two theories. For example, if we had data on insider ownership, or on the nature of the pay-for-performance relation for each of the banks in our sample, we could check to see if—as Brickley, Linck, and Smith implicitly suggest—these variables eliminate the effect of bank size in our regressions. Unfortunately, we do not have such data, so our ability to disentangle the two theories is limited.

3. Data

3.1. Sources

Our primary data source is the Federal Reserve's 1993 National Survey of Small Business Finance (NSSBF), which covers the financing practices of a stratified random sample of firms. The survey was actually conducted in 1994 and 1995 based on a sample of firms that were in existence at the end of 1993. Some of the information collected—e.g., on the firm's most recent loan—comes from the calendar year 1994. To be in the sample, a firm must be a for-profit entity with fewer than 500 employees. Consequently, the firms in our sample are really quite small, with a mean book value of assets of \$3.0 million, and a median of \$680,000.

The survey's focus on small firms is ideal for our purposes, for several reasons. First, many of the firms in our sample do not have formal financial records. This makes it plausible that soft information might play an important in evaluating their creditworthiness. Second, these firms secure most of their external finance from debt markets, and a predominant share of this comes from banks. Thus there is the possibility that being matched with the "wrong" kind of bank could have a meaningful effect on their overall access to finance. Third, with such small firms, the decision of whether to borrow from a large or small bank is unlikely be driven by regulatory lending limits.

Although the survey includes a complete inventory of all of a firm's current loans, we focus on its most recent loan, and only if that loan is from a bank. This allows us to focus on a fairly static banking environment, and also ensures that we measure firms' and banks' characteristics at roughly the time that loans are originated. In particular, each observation in our sample is based on a firm that secured a loan from its bank between 1990 and 1994; 88% of these loans were originated in either 1993 or 1994.

Each firm is then matched with the specific bank from which it borrows. For the banks, we use the Consolidated Report of Condition and Income (a.k.a. the Call Reports) to obtain balance-sheet variables such as bank assets. We also use the FDIC Summary of Deposits to determine the locations of individual bank branches. Our baseline sample includes 1,131 firms for which we have data on the most recent lender.

3.2. Variable definitions

In the analysis that follows, we work with the following basic variables. First, we have five variables that can be thought of as proxies for the nature of the relationship between the firm and its bank: (1) Distance is the number of miles between the firm and the bank branch or office from which the most recent loan was granted; (2) Impersonal Relationship is a dummy that equals one if the firm primarily communicates with the bank by phone or mail, and zero if the communication is face-to-face; (3) Relationship Length is the number of years that the bank has been providing services to the firm; (4) Single Lender is a dummy that equals one if the bank making the most recent loan is the firm's *only* (bank or non-bank) lender; and (5) Trade Credit Paid Late is the fraction of trade credit that the firm reports paying when it is past due. (The survey asks for the proportion of trade credit that is paid late on a scale from 1 to 5. For ease of interpretation, we recode this variable to be between zero and one.)

Next, there are six variables that capture bank and banking-market characteristics: (1) Bank Size is the assets of the firm's bank, expressed in billions of dollars; (2) Number of Branches in Market is the number of branches that the firm's bank has in the MSA or non-MSA rural county in which the firm is located; (3) Bank Age is the number of years the bank has been in existence; (4) Median Bank Size is the median assets across all banks (weighted by branches) in the firm's market; (5) Open Market is the fraction of the ten years prior to our sample period (i.e., 1983–1992) during which the firm's state was neither a unit banking nor limited branching state; and (6) Market Herfindahl is the banking-market Herfindahl index for this market (a measure of industry concentration). Bank Size will be the key right-hand-side variable of interest in most of our regressions, and both Median Bank Size and Open Market will be used as instruments for Bank Size.

Finally, there are eight variables that measure firm and contract characteristics: (1) Firm Size is the firm's assets, in millions of dollars; (2) Firm Age is the number of years the firm has been in existence; (3) Loan Amount is the size of the most recent loan, in millions; (4) Line of Credit is a dummy that takes on the value one if the most recent loan is a line of credit; (5) Loan Collateralized is a dummy that takes on the value one if the most recent loan is secured; (6) Checking Account is a dummy that takes on the value one if the firm also has a checking account with the bank that made its most recent loan; (7) Firm in MSA is a dummy that takes on the value one if the firm is located in an MSA; and (8) Records is a dummy that takes on the value one if the firm's respondent to the NSSBF survey said "yes" when asked if he or she had documentation such as financial statements or accounting records to help in answering the survey questions. (In what follows, we sometimes imprecisely refer to a "yes" answer to this question as indicating that a firm keeps formal records; it is more accurate to say that we only have a proxy for the existence of records—using records to answer the survey is likely to be highly correlated with having records in the first place, but it is not exactly the same thing.)

3.3. Summary statistics by bank size class

Table 1 presents summary statistics for many of the variables for both the full sample (in Panel A) and for subsamples based on bank size (in Panel B). Although

Table 1

Summary statistics Panel A: full sample

Panel A contains summary statistics for the variables used in all subsequent estimation. Distance is the distance between a firm and the bank branch or office it uses most often. Impersonal Relationship equals one if the firm interacts with its bank most often by phone or mail and zero if the interaction is in person. Relationship Length is the number of years the bank and the firm have been interacting (through lending, deposit, or service activities). Single Lender is a dummy variable that equals one if the firm has a single lender. Trade Credit Paid Late is the fraction of its trade credit the firm reports paying when it is past due. Bank Size is the assets of the bank from which the firm has its most recent loan. Number of Branches in Market is the number of branches which the *firm's bank* has in its market (MSA or county). Median Bank Size is the size of the median bank in the *firm's market* (MSA or county) weighted by branches. Open Market is the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state. Firm Size is the assets of the firm. Loan Amount is the size of the most recent loan. Records is a dummy variable that equals one if the person answering the income statement and balance sheet questions for the firm had documentation such as financial statements or accounting records to help answer the questions. There are 1,131 observations in the sample.

Variable	Mean	Std Dev	25%	50%	75%
Lending methods					
Distance (miles)	26.053	136.992	1.000	3.000	10.000
Impersonal relation (1 = yes)	0.294	0.456	0.000	0.000	1.000
Relationship length (yrs)	8.750	7.508	3.000	6.000	12.000
Single lender (1 = yes)	0.499	0.500	0.000	0.000	1.000
Trade credit paid late	0.352	0.208	0.250	0.250	0.500
Bank characteristics					
Bank size ($B)	8.883	23.147	0.163	0.956	7.685
No. of branches in market	21.486	45.494	1.000	5.000	25.000
Bank age (years)	75.263	43.914	39.000	80.000	106.000
Median bank size ($b)	6.159	13.426	0.196	1.203	6.077
Open market	0.446	0.266	0.000	0.400	0.800
Firm characteristics					
Firm age (years)	14.842	8.865	8.000	13.000	22.000
Firm size ($m)	3.003	7.136	0.150	0.680	2.850

Table 1 (*continued*)

Variable	Mean	Std Dev	25%	50%	75%
Loan amount ($m)	1.001	3.750	0.030	0.125	0.600
Records (1 = yes)	0.570	0.495	0.000	1.000	1.000

Panel B: means by bank size

Panel B contains the means of selected variables across four categories of bank size (less than $100M, $100M-1B, $1B-10B, and over $10B in assets). Regressions estimating how the lending method variables depend upon bank size as well as on other firm and bank characteristics are contained in later tables.

Variable	<100M	100M-1B	1B-10B	10B+
Lending methods				
Distance (miles)	14.947	9.488	19.302	71.363
Impersonal relation (1 = yes)	0.168	0.216	0.375	0.406
Relationship length (yrs)	9.384	9.261	8.762	7.389
Single lender (1 = yes)	0.616	0.496	0.497	0.410
Trade credit paid late	0.325	0.374	0.340	0.349
Bank characteristics				
Bank size ($b)	0.058	0.386	4.346	36.167
No. of branches in market	1.442	5.158	24.140	60.487
Bank age (years)	49.111	67.858	86.543	92.679
Median bank size ($b)	2.765	4.401	5.304	12.964
Open market	0.305	0.413	0.497	0.544
Firm characteristics				
Firm age (years)	13.763	15.037	15.595	14.346
Firm size ($m)	0.704	1.752	3.860	5.695
Loan amount ($m)	0.180	0.375	1.198	2.402
Records (1 = yes)	0.474	0.562	0.576	0.654
Number of observations	190	379	328	234

the firms in our sample are small (fewer than 500 employees), we still see a significant range of firm and loan sizes.[2] The range of bank sizes is even larger, increasing from $163 million in assets at the 25th percentile of the distribution to $7.69 billion in assets at the 75th percentile. Although these banks are selected because a small firm has borrowed from them, they are not exclusively small banks. In fact, they appear to be somewhat larger than is typical in a comprehensive sample of banks. For example, the 25th percentile of bank assets in our sample corresponds to roughly the 80th percentile of the size distribution of *all* banks in 1993 as reported in Kashyap and Stein (2000).

[2]The NSSBF does not use an equal-probability sample design but does include a weighting scheme that can be used to make the survey nationally representative. We choose not to employ the weights in the analysis presented here, since our hypotheses apply with equal force to all observations. However, we obtain very similar results if we run weighted versions of all our regressions.

As Panel B of Table 1 makes clear, there is a strong univariate correlation between bank size and many of the other variables. For example, mean loan size increases from \$180,000 in the smallest class of banks (those with assets below \$100 million) to \$2.40 million in the largest size class (those with assets above \$10 billion). Firm size increases similarly. The fraction of firms with financial records goes from 47.4% in the smallest class of banks to 65.4% in the largest class.

The aspects of lending relationships that we are interested in also vary across bank size classes in the manner predicted by the theory. The average distance between a firm and its bank rises from 14.9 miles for the smallest class of banks to 71.4 miles for the largest. Relatedly, the incidence of impersonal communication increases from 16.8% among the smallest banks to 40.6% among the largest banks. Mean relationship length is 9.4 years in the smallest class of banks, and 7.4 years in the largest class. The incidence of exclusive relationships is 61.6% among the smallest banks, and 41.0% among the largest banks.

3.4. The availability of small-bank branches in different markets

Our instrumenting strategy relies on being able to identify markets where there are substantial differences in the availability of small banks. To get a sense for the variation along this dimension, for each market in our sample, we calculate the fraction of branches that are owned by small banks, defined as those with assets of less than \$100 million. This variable has a mean of 31%. Not surprisingly, its correlation with our median-bank-size instrument is very strong, at −0.75. Moreover, when we regress the former on the latter, we find that increasing median bank size by two standard deviations decreases the fraction of small-bank branches in a market by 23 percentage points, which looms large relative to the mean of 31%. Thus, movements in the median-bank-size instrument would seem to translate into economically large changes in the availability of small-bank branches. Even more vividly, the 15 markets with the fewest small-bank branches have literally *no* branches owned by banks with less than \$100 million in assets, and hardly any (never more than 5%) branches owned by banks with less than \$300 million in assets. It is also worth noting that these markets are of widely varying sizes—they range from small markets like Burlington, VT to middle-sized ones like Lancaster, PA to larger population centers like Honolulu, HI, Jersey City, NJ, and Cleveland, OH.

4. Regression results

4.1. The choice of bank

We start by asking what determines the size of the bank from which a firm borrows. In Column 1 of Table 2, we use ordinary least squares (OLS) to regress Ln(Bank Size) against the firm and contract characteristics: Ln(Firm Size), Ln(1 + Firm Age), Ln(Loan Amount), Line of Credit, Loan Collateralized, Checking Account, Firm in MSA, and Records. The regression also includes

Table 2
Determinants of bank size

The dependent variable is Ln(Bank Size). Bank Size is expressed in thousands of dollars and Firm Size and Loan Size are expressed in dollars before taking logs. Ln(Median Bank Size) is the log of the median bank assets in the *firm's market* (MSA or county) weighted by branches. Open Market is the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state. We use Ln(Median Bank Size) and Open Market to instrument for Ln(Bank Size) in the models that follow. Each regression contains dummy variables for whether the loan is a line of credit, whether it is collateralized, and whether the firm has a checking account from the bank. Records is a dummy variable that equals one if the person answering the income statement and balance sheet questions for the firm had documentation such as financial statements or accounting records to help answer the questions. Each regression also includes dummies for the firm's industry (construction, retail, or services) and the year in which the loan was secured (1992–1994). The number of observations is 1,131. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Independent variables	Models	
	1: OLS	2: OLS
Bank and market characteristics		
Ln(median bank size)		0.222***
		(0.032)
Open market		0.438***
		(0.145)
Ln(1 + No. of branches)		0.594***
		(0.039)
Market Herfindahl		−0.226
		(0.541)
Ln(1 + bank age)		0.523***
		(0.051)
Firm and contract characteristics		
Ln(firm size)	0.125**	0.099**
	(0.051)	(0.040)
Ln(1 + firm age)	−0.172*	−0.224***
	(0.094)	(0.073)
Ln(loan amount)	0.277***	0.198***
	(0.051)	(0.040)
Line of credit (1 = yes)	0.197	0.110
	(0.141)	(0.110)
Loan collateralized (1 = yes)	−0.313**	−0.059
	(0.134)	(0.105)
Checking account (1 = yes)	−0.467**	−0.869***
	(0.210)	(0.166)
Firm in MSA (1 = yes)	1.220***	0.040
	(0.140)	(0.151)
Records (1 = yes)	0.240**	0.178*
	(0.121)	(0.094)
Adjusted R^2	0.218	0.526

dummies—not shown in the table—for the firm's industry (construction, retail, or services) as well as for the year in which the most recent loan was made.

As expected, bank size is strongly correlated with both the size of the firm in question and the size of the loan. If the size of the firm and the size of the loan both double, the regression tells us that bank assets increase by about 40%. But perhaps the most interesting result from this regression is the coefficient on Records, which is 0.240 and significant at the 5% level. Controlling for firm size, firms that have financial records borrow from banks that are roughly 24% larger. This is consistent with the idea that, all else equal, larger banks are at a comparative advantage in lending to firms for which hard information is more readily available.

One objection to this conclusion is that the Records variable could be endogenous. In particular, some firms might go to larger banks (for whatever reason) that then encourage record-keeping, so that the causality flows from bank size to records, rather than the other way around. Although it is hard to completely rule out this possibility—and given the story we have in mind, it is not clear that we would want to—we can make a partial attempt, as follows. If bank size drives record-keeping, then the correlation between bank size and records ought to be less pronounced for borrowers whose relationship to the bank is new, since the bank has not yet had time to really influence their behavior. However, this turns out not to be the case. We rerun the regression in Column 1 of Table 2, adding an interaction term given by the product of the Records variable and a dummy for whether the relationship is less than two years old (not shown). The interaction term is significantly positive, suggesting that, if anything, the correlation between bank size and records is *more* pronounced for newer relationships.

As discussed above, in our subsequent regressions we will use Ln(Bank Size) as an explanatory variable, and we will employ Ln(Median Bank Size) and Open Market as instruments for Ln(Bank Size). In Column 2 of Table 2, we display the first-stage regression that underlies this instrumenting procedure. In particular, we keep Ln(Bank Size) on the left, and add to the specification of Column 1 the following bank and banking-market variables: Ln(Median Bank Size), Open Market, Ln(1 + Number of Branches), Ln(1 + Bank Age), and Market Herfindahl. All of the right-hand-side variables in Column 2 of Table 2 will be controls in future regressions, except Ln(Median Bank Size) and Open Market, which will serve as the instruments for Ln(Bank Size). The main point to draw from this regression is that both Ln(Median Bank Size) and Open Market appear sufficiently correlated with Ln(Bank Size) to be viable instruments. They attract economically large coefficients and are highly statistically significant, with t-stats of 6.9 and 3.0, respectively.

4.2. *The distance between firms and their banks*

Table 3 examines the link between bank size and distance. In Column 1, we run an OLS regression in which the dependent variable is Ln(1 + Distance). The explanatory variables include the bank and banking-market characteristics Ln(Bank Size), Ln(1 + Number of Branches), Ln(1 + Bank Age), and Market Herfindahl, as well as the firm and contract characteristics Ln(Firm Size), Ln(1 + Firm Age),

Table 3
Distance between the firm and its bank

The dependent variable is the log of one plus the distance (in miles) between the firm and the bank branch or office that it uses most often. Ln(Bank Size) is the log of bank assets. Bank Size is expressed in thousands of dollars and Firm Size and Loan Size are expressed in dollars before taking logs. In Column 2, we report instrumental-variable estimates where the instruments for Ln(Bank Size) are Ln(Median Bank Size), the log of the median assets of banks in the area where the firm is located, and Open Market, the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state. The number of branches in the market includes only branches of the bank from which the firm borrows. Each regression contains dummy variables for whether the loan is a line of credit, whether it is collateralized, and whether the firm has a checking account from the bank. Records is a dummy variable that equals one if the person answering the income statement and balance sheet questions for the firm had documentation such as financial statements or accounting records to help answer the questions. Each regression also includes dummies for the firm's industry (construction, retail, or services) and the year in which the loan was secured (1992–1994). The number of observations is 1,131. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Independent variables	Models	
	1: OLS	2: IV
Bank and market characteristics		
Ln(bank size)	0.184***	0.296***
	(0.021)	(0.078)
Ln(1 + No. of branches)	−0.385***	−0.467***
	(0.031)	(0.063)
Market Herfindahl	−0.352	−0.455
	(0.392)	(0.403)
Ln(1 + bank age)	0.002	−0.049
	(0.038)	(0.052)
Firm and contract characteristics		
Ln(firm size)	0.028	0.018
	(0.030)	(0.031)
Ln(1 + firm age)	−0.216***	−0.189***
	(0.054)	(0.058)
Ln(loan amount)	0.076**	0.049
	(0.030)	(0.035)
Line of credit (1 = yes)	0.121	0.107
	(0.082)	(0.083)
Loan collateralized (1 = yes)	0.049	0.066
	(0.078)	(0.079)
Checking account (1 = yes)	−0.870***	−0.758***
	(0.125)	(0.147)
Firm in MSA (1 = yes)	0.255**	0.200*
	(0.105)	(0.112)
Records (1 = yes)	−0.026	−0.046
	(0.070)	(0.072)
Adjusted R^2	0.269	0.235

Ln(Loan Amount), Line of Credit, Loan Collateralized, Checking Account, Firm in MSA, and Records. In Column 2, we run the same basic regression by IV, using Ln(Median Bank Size) and Open Market as instruments for Ln(Bank Size). These regressions, and all those that follow, also include suppressed dummies for the firm's industry and the year the most recent loan was made.

Consistent with our theoretical prediction, firms that are customers of larger banks borrow at substantially greater distances. Both the OLS and the IV coefficients are statistically significant at the 1% level, and the IV coefficient is larger in magnitude, 0.296 versus 0.184. According to the IV estimate, increasing bank size from \$163 million in assets (the 25th percentile) to \$7.69 billion in assets (the 75th percentile) raises the predicted distance between a firm and its lender by 114%.

It is also worth briefly discussing some of the other controls in the regression and their importance. First, and not surprisingly, we find that the number of branches that the firm's lender has in the market is an important determinant of distance. Since larger banks naturally have more branches than small banks, it is especially important that we control for the number of branches in our tests. What the regression is then telling us is that the distance between a firm and its bank is positively related to *the size of the bank outside of the firm's local market*. In other words, if the bank adds branches outside of the firm's market, distance increases, but if the bank adds branches inside the firm's market, distance decreases, for the obvious mechanical reasons. We have verified this statement by re-running the basic OLS and IV regressions in Table 3, replacing Ln(Bank Size) with the log of one plus the number of branches that the bank has *outside* the market in question. In both cases, the coefficient on this variable is also positive and strongly significant.

We also find that older firms tend to be closer to their banks. At first, this seems puzzling because older firms might be expected to have better-established reputations (Diamond 1991), which should facilitate borrowing at a distance. The answer to the puzzle could be that firm age proxies for when the relationship was started. Petersen and Rajan (2002) and Hannan (2003) find that the distance between firms and their banks has been growing over time, partly because of the greater availability of hard information. So older firms could be closer to their banks because they started their relationships at a time when little hard public information was available about them. Finally, firms that have checking accounts with their banks are also closer to them. This replicates a finding in Petersen and Rajan (2002), and may be explained by the greater necessity of making physical trips to the bank with which one has a checking account.

A couple of other points deserve mention. The literature on bank consolidation has raised the question of whether banking mergers disrupt relationships, especially those that rely on soft information. Thus, when we find that larger banks are more likely to lend at a distance, we want to be sure that our bank size result is not due only to the effect of mergers. To test this, we re-run our basic specification, adding two controls for bank mergers (in regressions not reported in the tables). These variables are individually insignificant and make no material difference to our principal conclusions. In a similar spirit, we also add more controls for bank health and firm risk; again our results are qualitatively unaffected.[3]

4.3. The mode of conducting business: personal vs. impersonal

In Table 4, we investigate the link between bank size and the mode of communication. The right-hand-side variables are exactly the same as in Table 3, and the left-hand-side variable is now Impersonal Relationship. Also, given the dichotomous nature of the Impersonal Relationship variable, we run the regressions by logit, instead of by OLS (though our results are virtually identical if we use OLS). In Column 1, we use Ln(Bank Size) directly in the logit regression, and in Column 2 we instrument for Ln(Bank Size) by replacing it with its fitted value from the first-stage regression in Column 2 of Table 2.

Both the ordinary logit and the IV version yield strong, statistically significant estimates for the influence of bank size on the mode of communication, though as before the IV estimate is noticeably bigger. Based on the IV coefficient, an increase in bank size from the 25th to the 75th percentile raises the probability of impersonal communication from 15% to 38%. As with distance, the number of branches that the firm's bank has in the local market also affects the way in which the firm and the bank interact. In this case, having more in-market branches leads to significantly less impersonal communication, as would be expected.

Impersonal communication and physical distance are clearly related—it is more difficult to visit a distant bank in person. As a more stringent test, we can ask if bank size affects the mode of communication even after controlling for distance. In an unreported regression, we find that when we add Ln(1 + Distance) to the right-hand side of the IV specification, the coefficient on Ln(Bank Size) drops from 0.324 to 0.160. Although this still represents an economically interesting magnitude, the point estimate is no longer statistically significant. In the un-instrumented logit specification, adding Ln(1 + Distance) drops the coefficient on Ln(Bank Size) from 0.196 to 0.096, but in this case the coefficient remains significant at the 10% level.

With respect to firm characteristics, we find strong evidence that larger firms are more likely to communicate impersonally with their bankers, which is not surprising. At the same time, controlling for size, we find that older firms are *less likely* to communicate impersonally. This is at least in part driven by the earlier finding that older firms are physically closer to their banks. There may also be a vintage effect at work, whereby managers of older firms started off their careers interacting with their bankers face-to-face and have not changed their ways, even as the technology of

[3]As added bank controls, we include dummy variables for each of the following: whether a bank was the surviving bank in a merger in the last three years; whether the bank changed top-tier holding companies in the last three years; whether the bank's equity to asset ratio was in the bottom 10% of our sample; and whether the bank's ratio of non-performing loans to all loans was in the top 10% of our sample. As added firm-risk controls, we include the firm's leverage ratio and dummies for whether, over the past three years, any of the following have occurred: the owner has been 60 or more days delinquent on personal obligations; the firm has been 60 or more days delinquent; or a judgment was rendered against the owner.

Table 4
Impersonal communication between the firm and its bank

The dependent variable is a dummy that equals one if the bank and firm communicate impersonally (by phone or mail) and zero if they communicate in person. A logit model was estimated. Ln(Bank Size) is the log of bank assets. Bank Size is expressed in thousands of dollars and Firm Size and Loan Size are expressed in dollars before taking logs. In Column 2, we report instrumental-variable estimates where Ln(Bank Size) is replaced with its predicted value based on Ln(Median Bank Size), the log of the median assets of banks in the area where the firm is located, and Open Market, the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state (see Table 2, Column 2). The number of branches in the market includes only branches of the bank from which the firm borrows. Each regression contains dummy variables for whether the loan is a line of credit, whether it is collateralized, and whether the firm has a checking account from the bank. Records is a dummy variable that equals one if the person answering the income statement and balance sheet questions for the firm had documentation such as financial statements or accounting records to help answer the questions. Each regression also includes dummies for the firm's industry (construction, retail, or services) and the year in which the loan was secured (1992–1994). The number of observations is 1,131. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Independent variables	Models	
	1: logit	2: logit/IV
Bank and market characteristics		
Ln(bank size)	0.196***	0.324**
	(0.046)	(0.165)
Ln(1 + No. of branches)	−0.267***	−0.365***
	(0.064)	(0.131)
Market Herfindahl	−0.808	−0.883
	(0.907)	(0.916)
Ln(1 + bank age)	−0.132	−0.179
	(0.082)	(0.108)
Firm and contract characteristics		
Ln(firm size)	0.259***	0.248***
	(0.070)	(0.070)
Ln(1 + firm age)	−0.329***	−0.290**
	(0.119)	(0.124)
Ln(loan amount)	0.082	0.049
	(0.066)	(0.075)
Line of credit (1 = yes)	0.659***	0.642***
	(0.191)	(0.190)
Loan collateralized (1 = yes)	0.215	0.245
	(0.170)	(0.170)
Checking account (1 = yes)	−1.128***	−0.982***
	(0.266)	(0.302)
Firm in MSA (1 = yes)	0.687***	0.608**
	(0.245)	(0.255)
Records (1 = yes)	−0.109	−0.115
	(0.153)	(0.154)
Pseudo R^2	0.179	0.168

banking has evolved. We also find that firms that have checking accounts with the bank are more inclined to meet with their banker face-to-face.

4.4. The effect of bank size on relationship length and exclusivity

Table 5 looks at the effect of bank size on relationship length. The structure is identical to that of Table 3, except that the dependent variable is now Ln(1 + Relationship Length). Relationships are significantly shorter when the firm borrows from a larger bank. According to the IV specification in Column 2 of Table 5, an increase in bank size from the 25th to the 75th percentile cuts the predicted length of a relationship almost in half, shrinking it from 8.8 to 4.5 years.

It should be noted that the estimated coefficient on Ln(Bank Size) is nearly three times higher in Column 2, where we use IV, as compared to Column 1, where we use OLS. Indeed, the theory suggests that it ought to be particularly important to deal with the endogeneity of bank size here, since firms might be more prone to switch to small banks—thereby setting the relationship-length clock back to zero—if they get into trouble and become the sort of "difficult" credits for which soft information is especially important. This would obviously make it hard to find an OLS association between small banks and longstanding relationships.

Table 6 analyzes the exclusivity of banking relationships, putting Single Lender on the left-hand side of the regressions. As in Table 4, the regressions are run with logit, given the dichotomous nature of the dependent variable. According to the IV specification, the effect of bank size on exclusivity is extremely strong: an increase in bank size from the 25th to the 75th percentile reduces the probability of an exclusive relationship by almost 50 percentage points, from 74% to 27%.

Again, we see the importance of instrumenting, as the coefficient on Ln(Bank Size) goes from –0.096 in the ordinary logit to –0.526 in the instrumented version. And again, this makes perfect sense in light of the theory. Petersen and Rajan (1994) show that troubled firms are more likely to have multiple relationships—presumably as they cast around for someone willing to accommodate their needs—and our theory suggests that troubled firms should also be more prone to match with small banks. Hence we would expect the non-instrumented coefficient to be significantly biased towards zero.

At this juncture, it might be useful to ask whether the effects of bank size on distance and on the mode of interaction work only indirectly through the kind of relationship (long and exclusive with small banks, short and non-exclusive with large banks), or whether there is also a direct effect. One way to test this is to include both Ln(1 + Relationship Length) and Single Lender as additional controls in the regressions of Tables 3 and 4, where the dependent variables are Ln(1 + Distance) and Impersonal Relationship, respectively. In both cases, the coefficients on Ln(Bank Size) continue to be strongly statistically significant and only slightly diminished in magnitude, suggesting that bank size indeed has an important independent effect.

Table 5
Relationship length between the firm and its bank

The dependent variable is log of one plus the length of the relationship between the firm and its bank (in years). Ln(Bank Size) is the log of bank assets. Bank Size is expressed in thousands of dollars and Firm Size and Loan Size are expressed in dollars before taking logs. In Column 2, we report instrumental-variable estimates where the instruments for Ln(Bank Size) are Ln(Median Bank Size), the log of the median assets of banks in the area where the firm is located, and Open Market, the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state. The number of branches in the market includes only branches of the bank from which the firm borrows. Each regression contains dummy variables for whether the loan is a line of credit, whether it is collateralized, and whether the firm has a checking account from the bank. Records is a dummy variable that equals one if the person answering the income statement and balance sheet questions for the firm had documentation such as financial statements or accounting records to help answer the questions. Each regression also includes dummies for the firm's industry (construction, retail, or services) and the year in which the loan was secured (1992–1994). The number of observations is 1,131. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Independent variables	Models	
	1: OLS	2: IV
Bank and market characteristics		
Ln(bank size)	−0.048***	−0.150***
	(0.012)	(0.044)
Ln(1 + No. of branches)	0.051***	0.125**
	(0.017)	(0.035)
Market Herfindahl	0.313	0.408*
	(0.215)	(0.225)
Ln(1 + bank age)	0.108***	0.155***
	(0.021)	(0.029)
Firm and contract characteristics		
Ln(firm size)	0.022	0.031*
	(0.016)	(0.017)
Ln(1 + firm age)	0.607***	0.582***
	(0.030)	(0.032)
Ln(loan amount)	−0.045***	−0.020
	(0.016)	(0.020)
Line of credit (1 = yes)	0.000	0.012
	(0.045)	(0.046)
Loan collateralized (1 = yes)	−0.056	−0.072
	(0.043)	(0.044)
Checking account (1 = yes)	0.446***	0.343***
	(0.068)	(0.082)
Firm in MSA (1 = yes)	−0.165***	−0.115**
	(0.057)	(0.063)
Records (1 = yes)	−0.049	−0.031
	(0.038)	(0.040)
Adjusted R^2	0.366	0.348

Table 6
Exclusive relationship between the firm and its bank

The dependent variable is a dummy that equals one if the bank is the firm's only lender, and zero otherwise. A logit model is estimated. Ln(Bank Size) is the log of bank assets. Bank Size is expressed in thousands of dollars and Firm Size and Loan Size are expressed in dollars before taking logs. In Column 2, we report instrumental-variable estimates where Ln(Bank Size) is replaced with its predicted value based on Ln(Median Bank Size), the log of the median assets of banks in the area where the firm is located, and Open Market, the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state (see Table 2, Column 2). The number of branches in the market includes only branches of the bank from which the firm borrows. Each regression contains dummy variables for whether the loan is a line of credit, whether it is collateralized, and whether the firm has a checking account from the bank. Records is a dummy variable that equals one if the person answering the income statement and balance sheet questions for the firm had documentation such as financial statements or accounting records to help answer the questions. Each regression also includes dummies for the firm's industry (construction, retail, or services) and the year in which the loan was secured (1992–1994). The number of observations is 1,131. Significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

Independent variables	Models	
	1: logit	2: logit/IV
Bank and market characteristics		
Ln(bank size)	−0.096**	−0.526***
	(0.040)	(0.144)
Ln(1 + No. of branches)	0.075	0.388***
	(0.057)	(0.116)
Market Herfindahl	−0.637	−0.242
	(0.721)	(0.732)
Ln(1 + bank age)	0.115	0.308**
	(0.071)	(0.095)
Firm and contract characteristics		
Ln(firm size)	−0.318***	−0.283***
	(0.057)	(0.058)
Ln(1 + firm age)	0.246**	0.143
	(0.102)	(0.107)
Ln(loan amount)	0.112**	0.219***
	(0.056)	(0.065)
Line of credit (1 = yes)	−0.078	−0.027
	(0.150)	(0.152)
Loan collateralized (1 = yes)	−0.351**	−0.149***
	(0.142)	(0.145)
Checking account (1 = yes)	0.517**	0.087
	(0.235)	(0.271)
Firm in MSA (1 = yes)	−0.098	0.110
	(0.192)	(0.204)
Records (1 = yes)	−0.120	−0.044
	(0.128)	(0.131)
Pseudo R^2	0.062	0.067

4.5. Bank size and credit availability

Thus far, we have argued that soft information is likely to be important in evaluating the creditworthiness of small firms, and that small banks have a comparative advantage in acquiring and acting on such soft information, which is why they can form stronger relationships with the firms in our sample. But do these stronger relationships translate into more financing? In other words, do they have meaningful real effects?

The problem in measuring the availability of credit is that we cannot simply look at the amount of debt on a firm's balance sheet, for that will reflect both demand and supply considerations. But we can use an alternative approach, following Petersen and Rajan (1994). The idea is that if banks (or any other intermediaries) limit the credit extended to a firm, the firm will be forced to borrow from a more expensive source. Holding investment opportunities constant, the amount borrowed from the more expensive sources should measure the degree of credit rationing by banks.

Petersen and Rajan (1994, 1995) point to stretched trade credit as an extremely costly marginal source of finance, and argue that the fractional share of a firm's trade credit that is paid late provides a reliable measure of the extent to which the firm is rationed. Older and larger firms, which are presumably less constrained by banks, pay less of their trade credit late. Similarly, firms that have long-term relationships with their banks also pay less of their trade credit late.

In Table 7, we repeat our basic specification, putting Trade Credit Paid Late on the left-hand side. Given that this variable is bounded between zero and one, we run the regressions with a two-sided Tobit procedure. (Again, however, we get essentially identical results—both with and without instrumenting—if we use ordinary least squares instead of Tobit.) It should also be noted that the number of observations in Table 7 is reduced from 1,131 to 546, because we do not have the trade-credit data for all of the firms in our sample.

It is in these regressions that instrumenting for bank size is most important. The coefficient on Ln(Bank Size) is small and statistically insignificant in Column 1, when we enter it directly in the regression. But when we instrument for it in Column 2 with Ln(Median Bank Size), the coefficient becomes statistically significant and economically large. In particular, the IV estimate implies that an increase in bank size from the 25th percentile to the 75th percentile raises the fraction of trade credit that is paid late by 17 percentage points, from 26% to 43%. The bottom line is that firms that are forced to borrow from large banks appear to be substantially more credit constrained than those that can borrow from small banks.

When we test formally whether bank size is exogenous in this model, we reject the hypothesis, with a p-value of 0.06. This is seen in the Hausman (1978) test in Column 3 of Table 7. The sign of the bias, however, is again interesting. The endogenous portion of bank size (i.e., the residual from the first-stage regression in Column 2 of Table 2) is *negatively* correlated with Trade Credit Paid Late. To the extent that they can choose, firms that are more prone to being credit rationed pair up with smaller banks. This endogenous pattern of firm-bank matching fits with both the theory and all of the other evidence that we have documented so far. Given that small banks are

Table 7
Fraction of trade credit paid late

The dependent variable is the fraction of trade credit the firm pays late. A tobit model is estimated. Ln(Bank Size) is the log of bank assets. Bank Size is expressed in thousands of dollars and Firm Size and Loan Size are expressed in dollars before taking logs. In Column 2, we report instrumental-variable estimates where Ln(Bank Size) is replaced with its predicted value based on Ln(Median Bank Size), the log of the median assets of banks in the area where the firm is located, and Open Market, the fraction of the previous ten years during which there were no restrictions on within-state branching in the firm's state (see Table 2, Column 2). The number of branches in the market includes only branches of the bank from which the firm borrows. Each regression contains dummy variables for whether the loan is a line of credit, whether it is collateralized, and whether the firm has a checking account from the bank. Records is a dummy variable that equals one if the respondent to the survey had documentation to help answer the questions. Bank Size Residual is the residual from the first-stage bank-size regression (Table 2, Column 2) and is used to conduct a test of whether bank size is exogenous. Each regression also includes dummies for the firm's industry (construction, retail, or services) and the year in which the loan was secured (1992–1994). The number of observations is 546. Significance at the 10%, 5%, and 1% levels is denoted by $*$, $**$, and $***$, respectively.

Independent variables	Models		
	1: tobit	2: tobit/IV	3: tobit
Bank and market characteristics			
Ln(bank size)	0.006	0.044**	0.044**
	(0.006)	(0.021)	(0.021)
Ln(1 + No. of branches)	−0.017**	−0.044***	−0.044***
	(0.008)	(0.017)	(0.017)
Market Herfindahl	0.024	−0.016	−0.016
	(0.112)	(0.114)	(0.114)
Ln(1 + bank age)	0.007	−0.010	−0.010
	(0.010)	(0.014)	(0.014)
Bank size residual (Hausman test)			−0.041*
			(0.023)
Firm and contract characteristics			
Ln(firm size)	−0.002	−0.005	−0.005
	(0.009)	(0.009)	(0.009)
Ln(1 + firm age)	0.022	0.032**	0.032**
	(0.015)	(0.016)	(0.016)
Ln(loan amount)	−0.008	−0.017*	−0.017*
	(0.009)	(0.010)	(0.010)
Line of credit (1 = yes)	0.003	0.000	0.000
	(0.023)	(0.023)	(0.023)
Loan collateralized (1 = yes)	0.042*	0.050**	0.050**
	(0.022)	(0.022)	(0.022)
Checking account (1 = yes)	−0.003	0.036	0.036
	(0.039)	(0.045)	(0.045)
Firm in MSA (1 = yes)	0.066**	0.047	0.047
	(0.030)	(0.031)	(0.031)
Records (1 = yes)	0.028	0.022	0.022
	(0.020)	(0.020)	(0.020)
Log likelihood	12.425	14.090	14.180

better at building relationships based on soft personal information, we should expect firms that are having a hard time raising finance to be especially likely to turn to small banks for help.

Our results for trade credit are especially noteworthy, in the following sense. Some of our other findings—particularly those for distance and the mode of communication—can be seen as consistent with the hypothesis that, because they can better spread fixed costs, big banks use a strictly dominant technology (credit scoring), which in turn gives them an absolute advantage over smaller banks when it comes to small-business lending. But this hypothesis implies that big banks should do *better* at the task of relaxing credit constraints, which is the opposite of what we see in the data.

4.6. Banks vs. bank holding companies: whose size matters?

Any bank in our sample can be either a stand-alone bank, or part of a multi-bank holding company. The measure of bank size that we have been using throughout *does not* include the assets of other banks that are part of the same holding company. Moreover, 65% of our sample firms borrow from banks that are part of holding companies.

To examine the effects of holding-company structure, we try including an additional variable in all of our regressions: the log of assets of the *other* banks in the multi-bank holding company, if any exist. (This variation is not reported in the tables.) Interestingly, we find that, keeping the assets of the firm's own bank constant, the size of the rest of the holding company does not have a meaningful effect on any of the outcomes—i.e., distance, the mode of communication, the length and exclusivity of relationships, and the extent of credit constraints. Even in the few specifications where the size of the rest of the holding company yields a statistically significant coefficient, this coefficient is an order of magnitude smaller than that for the size of the firm's own bank. Similar results emerge if we instead add to our baseline specifications either: (i) a multi-bank holding-company dummy; or (ii) this dummy along with its interaction with the size of the firm's bank. In all cases, the coefficients on own bank size remain essentially unchanged, and the holding-company variables contribute little additional explanatory power.

These patterns can be thought of as supportive of Stein (2002). As argued in Section 2, Stein's control-rights-based model suggests that if decisions within the holding company can be credibly decentralized to the bank level, then the size of the holding company outside of the specific bank in question should not matter much. The evidence can also be rationalized within the explicit-incentives framework of Brickley et al. (2003). At first glance, one might interpret their theory as predicting that it is primarily the size of the holding company that ought to matter, since share ownership by individual loan officers is likely to be negatively related to holding-company size. However, there are other reasons why explicit incentives might be a function of the size of the bank rather than of the holding company. For example, individual bank subsidiaries within holding companies produce their own audited financial statements, which, as described by Blackwell et al. (1994), can facilitate the provision of a variety of performance-based incentives at the bank level (e.g., profit-linked bonuses, promotions, or terminations). And given the existence of such bank-level incentives, standard free-riding arguments might suggest that they would optimally be higher-powered when the bank in question is smaller.

While the distinction between the size of the bank vs. the size of the holding company might not allow us to cleanly separate the theories of Stein (2002) and Brickley et al. (2003), it is more helpful in cutting against other interpretations of some of our results. As noted above, one might argue that our findings for variables like distance simply reflect an *absolute advantage* of big banks. Perhaps credit scoring—and the associated tendency to lend at a greater distance—is just a better technology for lending, but only big banks can afford to use it because it involves substantial fixed costs. However, if this were the case, we would expect the size of the holding company to be the most significant determinant of distance, since the cost of a credit-scoring system can presumably be spread across an entire holding company.

4.7. Robustness: instrumenting with only the state-level regulatory variable

As noted above, it is possible to question the validity of one of our instruments, Ln(Median Bank Size): one can hypothesize that some markets have certain attributes that tend to attract both banks of a certain size and firms with particular characteristics. (One of the most obvious such attributes is population; however, when we replace the MSA dummy in all of our regressions with a log-population control, nothing changes.) So as an alternative, we try dropping Ln(Median Bank Size) and using the state-level regulatory variable Open Market as our only instrument.

On the one hand, Open Market is sufficiently correlated with Ln(Bank Size) that it would appear to be a workable instrument on its own—the univariate correlation between the two variables is 0.227. On the other hand, it is a weaker instrument than Ln(Median Bank Size), which has a correlation of 0.490 with Ln(Bank Size). Thus, this approach, while more conservative, also sacrifices considerable power.

Comfortingly, the results using Open Market as the only instrument are generally very close to those obtained with the two instruments together.[4] Moreover, if we adopt the identifying assumption that Open Market is exogenous, we can for each of our left-hand-side variables conduct a specification test of the hypothesis that Ln(Median Bank Size) is exogenous as well. This hypothesis is never rejected, which lends more support to the notion that Ln(Median Bank Size) is a legitimate instrument.

[4]In the Distance regression, we report in Table 3 an IV coefficient of 0.296 on Ln(Bank Size); this coefficient changes to 0.362 when we use Open Market as our only instrument. With Impersonal Relationship, the IV coefficient on Ln(Bank Size) goes from 0.324 to 0.207, with Relationship Length it goes from –0.150 to –0.189, with Single Lender it goes from –0.526 to –0.485, and with Trade Credit Paid Late it goes from 0.044 to 0.028. In spite of the increased standard errors, the estimates for Distance, Relationship Length, and Single Lender continue to be highly statistically significant (with *p*-values of 0.002, 0.002, and 0.017 respectively). The estimates for Impersonal Relationship and Trade Credit Paid Late, however, are no longer significant.

5. Connection to the banking literature

There is a large literature on banks' lending practices. Although we cannot provide a full survey of this work, we can sketch some of its broad contours, in an effort to show how our findings fit in. A first category of research has employed regulatory data on banks (such as the Call Reports and the Summary of Deposits used in this paper), *without* being able to match these data to information on the small businesses doing the borrowing. These studies typically find that large banks allocate far lower proportions of their assets to small-business loans than do small banks (e.g., Berger et al., 1995), and that ratios of small-business loans to assets tend to decline after large banks are involved in mergers and acquisitions (e.g., Peek and Rosengren, 1998; Strahan and Weston, 1996, 1998; Berger et al., 1998; Sapienza, 2002). Brickley et al. (2003) document that rural banking offices are more likely to be owned by small banks. Although small-business loans and rural locations are potentially correlated with soft information, these papers do not establish the link between bank size and soft information directly.

A second category of work has examined data on small businesses (such as the NSSBF survey that we use) but, again, without being able to match these data to information on the banks doing the lending. These studies find that stronger bank-borrower relationships are generally associated with better treatment for borrowers, in terms of lower interest rates and reduced collateral requirements (Berger and Udell, 1995), increased credit availability (Petersen and Rajan, 1994, 1995; Cole, 1998), and greater protection against interest rate shocks (Berlin and Mester, 1998). While all of these results help make the case that the soft information embedded in a banking relationship is valuable, none of them speak to the question of what kind of bank is best able to generate and act on soft information.

Finally, a handful of studies have used regulatory data on banks that *are* matched to their small-business borrowers, as in this paper. It has been found that large banks more often lend to larger, older, more financially secure firms (Haynes et al., 1999), and to firms that borrow from multiple banks (Berger et al., 2001).[5] Also, large banks charge relatively low interest rates and have low collateral requirements for small-business loans (Berger and Udell, 1996; Berger et al., 2003). All of these pieces of evidence fit with the idea that, within the general class of small-business loans, large banks systematically try to pick off the largest, safest, and easiest-to-evaluate credits. But it seems fair to say that none of them gets at the underlying mechanism that creates this pattern of behavior.

To our knowledge, only one previous paper has tried to directly examine how lending practices themselves differ between large and small banks. Cole et al. (2004) use survey data to look at the loan approval process across banks of different sizes. They find that for large banks (over \$1 billion in assets), approvals are based primarily on standard criteria obtained from financial statements—a so-called "cookie cutter" approach. In contrast, hard financial numbers have less explanatory

[5] Ongena and Smith (2001) find that large banks have longer relationships with firms, in contrast to our results. However, they use data on larger public firms that have access to other forms of external finance.

power (in an R^2 sense) for the approval decisions of small banks. This is consistent with small banks basing their decisions more heavily on soft information. It also ties in nicely with our results, which—though they are based on an entirely different set of variables and econometric procedures—suggest the same basic conclusions.

Black and Strahan (2002) find that the rate of creation of new businesses goes up in U.S. states that liberalize their banking laws, even while the share of small banks declines. While one might interpret this finding as evidence that small banks are not "special," there are other effects of liberalization. For example, competition increases, which is likely to increase both the number and efficiency of banks of all sizes, and hence to be beneficial for any given size distribution of banks. So it is hard to draw sharp conclusions from their work about whether small banks are particularly good at small-business lending. And conversely, none of our results should be taken to imply that the net effect of liberalization is negative. A similar point can be made about Jayaratne and Wolken (1999), who find that small firms are not more credit constrained in markets where large banks have a bigger fractional share—such markets could simply have a greater overall availability of nearby bank branches of any kind.

Finally, it is worth mentioning a recent paper that examines some of the same themes in the context of a different type of financial intermediation. Chen et al. (2005) show that small mutual funds tend to outperform large ones, particularly when it comes to investing in small-cap stocks that are located close by—i.e., stocks for which soft-information acquisition is likely to be most relevant. And in a striking parallel with our work, they find that it is the size of the *fund*, and not the size of the *fund family*, that matters for performance. Indeed, all else equal, funds belonging to large families actually perform better than those belonging to small families. As with our results for banks vs. holding companies, this would seem to suggest that decentralization of decision-making authority can have meaningful real effects.

6. Conclusions

While there has been much theoretical work by economists on the Coasian topic of organizations and their boundaries, there has been far less empirical work. A particularly under-explored set of empirical issues has to do with the ways in which an organization's form affects its ability to carry out different types of functions. The goal of this paper has been to take some first steps towards addressing these issues.

Our analysis is based on the premise that in small organizations, the center of decision-making authority is likely to be close to the point of information collection. According to Stein (2002), this creates strong incentives for soft-information production in small organizations. In contrast, large organizations have a tougher time providing incentives for their employees to produce soft information, although they tend to do well with respect to the creation of hard information. Large organizations also benefit from having broader internal capital markets—i.e., conditional on having acquired some hard information, they have more scope for actively reallocating resources based on this information. The bottom line is that one might expect small organizations to have a comparative advantage over large ones in activities that require the processing of a lot of soft information, and for the reverse to be true in activities that rely mostly on hard information.

In an effort to test this theory, we examine how banks of different sizes approach the task of small-business lending. We find that large banks lend primarily to larger firms with good accounting records, while small banks lend to more difficult credits. We also find that correcting for the endogeneity of the bank-firm match, large banks lend at a greater distance, interact more impersonally with their borrowers, have shorter and less exclusive relationships, and are not as effective at alleviating credit constraints. These effects are both statistically significant and economically large in magnitude, and they are all consistent with the hypothesis that small banks have a comparative advantage in lending based on soft information.

From a policy perspective, our results suggest that bank consolidation may raise meaningful concerns for small firms. Moreover, the key issue might be not so much about banks having market power in the traditional Herfindahl-index sense but rather, the degree to which firms have choice over the size of the bank they do business with. For it is when they have no choice and therefore have to borrow from large banks that our sample firms appear most prone to being credit constrained.

A similar policy-related observation can be made about the appeal to developing countries of encouraging entry by large multinational banks. Having foreign banking giants set up shop in a developing economy no doubt has a number of significant benefits. For example, they are probably more likely to be stable and financially sound. They might also be less likely to engage in the sort of corrupt related-lending practices documented by LaPorta et al. (2003). Without denying the importance of these factors, our analysis points to a potential tradeoff. If large foreign banks substantially crowd out smaller domestic ones, the supply of loans to informationally opaque small businesses could be negatively affected.

Finally, our results suggest that the standard practice in many countries of setting up large bureaucratic organizations to provide subsidized credit to small businesses (or alternatively, of forcing large banks to do so), might not be very effective. It might make more sense to target subsidies through smaller financial intermediaries, who can better incorporate soft information into their credit decisions.

While our analysis has focused on the banking industry, there are reasons to believe that the conclusions could generalize to a variety of other settings. Small-business lending is not unique in its reliance on soft information. Other relationship-based activities such as investment banking, consulting, and law also make heavy use of soft information. So too do certain kinds of research and new product development. Even some governmental activities, such as law enforcement, may require the creation and efficient use of substantial amounts of soft information. Our results suggest that, in all of these cases, organizational structure might play a crucial role in determining how effectively the job at hand is carried out. It would be instructive to study some of these other activities in detail, to see if this hypothesis is borne out more broadly in the data.

We have also found preliminary evidence—from the holding-company-level data—that seems to indicate that credible decentralization of decision-making can offset the effects of raw organizational size. This raises the possibility that a large organization might, at least to a degree, be able to enjoy the best of both worlds if it sets up an internal structure that achieves the right level of decentralization. Again, this is a conjecture that would greatly benefit from further empirical investigation.

References

Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. Journal of Political Economy 105, 1–29.

Baker, G.P., Hubbard, T., 2003. Make versus buy in trucking: asset ownership, job design and information. American Economic Review 93, 551–572.

Berger, A.N., Udell, G.F., 1995. Relationship lending and lines of credit in small firm finance. Journal of Business 68, 351–382.

Berger, A.N., Udell, G.F., 1996. Universal banking and the future of small business lending. In: Saunders, A., Walter, I. (Eds.), Financial System Design: The Case for Universal Banking. Irwin Burr, Ridge, IL, pp. 559–627.

Berger, A.N., Udell, G.F., 2002. Small business credit availability and relationship lending: the importance of bank organizational structure. Economic Journal 112, F32–F53.

Berger, A.N., Kashyap, A.K., Scalise, J.M., 1995. The transformation of the U.S. banking industry: what a long, strange trip it's been. Brookings Papers on Economic Activity, 55–218.

Berger, A.N., Saunders, A., Scalise, J.M., Udell, G.F., 1998. The effects of bank mergers and acquisitions on small business lending. Journal of Financial Economics 50, 187–229.

Berger, A.N., Demsetz, R.S., Strahan, P.E., 1999. The consolidation of the financial services industry: causes, consequences, and implications for the future. Journal of Banking and Finance 23, 135–194.

Berger, A.N., Klapper, L.F., Udell, G.F., 2001. The ability of banks to lend to informationally opaque small businesses. Journal of Banking and Finance 25, 2127–2167.

Berger, A.N., Rosen, R., Udell, G.F., 2003. Does market size structure affect competition? The case of small business lending. Unpublished working paper, Federal Reserve Board.

Berlin, M., Mester, L., 1998. On the profitability and cost of relationship lending. Journal of Banking and Finance 22, 873–897.

Black, S.E., Strahan, P.E., 2002. Entrepreneurship and bank credit availability. Journal of Finance 57, 2807–2833.

Blackwell, D.W., Brickley, J.A., Weisbach, M.S., 1994. Accounting information and internal performance evaluation: evidence from Texas banks. Journal of Accounting and Economics 17, 331–358.

Brickley, J.A., Linck, J.S., Smith, C.W., 2003. Boundaries of the firm: evidence from the banking industry. Journal of Financial Economics 70, 351–383.

Chen, J., Hong, H., Huang M., Kubik, J.D., 2005. Does fund size erode mutual fund performance? The role of liquidity and organization. American Economic Review, forthcoming.

Coase, R.H., 1937. The nature of the firm. Economica 4, 386–405.

Cole, R.A., 1998. The importance of relationships to the availability of credit. Journal of Banking and Finance 22, 959–977.

Cole, R.A., Goldberg, L.G., White, L.J., 2004. Cookie-cutter vs. character: the microstructure of small business lending by large and small banks. Journal of Financial and Quantitative Analysis 39.

Coval, J., Moskowitz, T., 2001. The geography of investment: informed trading and asset prices. Journal of Political Economy 109, 811–841.

DeYoung, R., Goldberg, L.G., White, L.J., 1998. Youth, adolescence and maturity of banks: credit availability to small business in an era of banking consolidation. Journal of Banking and Finance 23, 463–492.

Diamond, D.W., 1991. Monitoring and reputation: the choice between bank loans and directly placed debt. Journal of Political Economy 99, 689–721.

Fisman, R.J., Love, I., 2003. Trade credit, financial intermediary development and industry growth. Journal of Finance 58, 353–374.

Grossman, S.J., Hart, O.D., 1986. The costs and benefits of ownership: a theory of vertical and lateral integration. Journal of Political Economy 94, 691–719.
Hannan, T.H., 2003. Changes in non-local lending to small business. Journal of Financial Services Research 24, 31–46.
Hart, O.D., 1995. Firms, Contracts, and Financial Structure. Oxford University Press, Oxford.
Hart, O.D., Moore, J., 1990. Property rights and the nature of the firm. Journal of Political Economy 98, 1119–1158.
Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.
Haynes, G.W., Ou, C., Berney, R., 1999. Small business borrowing from large and small banks. In: Blanton, J.L., Williams, A., Rhine, S.L.W. (Eds.), Business Access to Capital and Credit. Federal Reserve System Research Conference, pp. 287–327.
Holmstrom, B., 1999. The firm as a subeconomy. Journal of Law Economics and Organization 15, 74–102.
Holmstrom, B., Milgrom, P., 1994. The firm as an incentive system. American Economic Review 84, 972–991.
Jayaratne, J., Wolken, J., 1999. How important are small banks to small business lending? New evidence from a survey of small firms. Journal of Banking and Finance 23, 427–458.
Kashyap, A.K., Stein, J.C., 2000. What do a million observations on banks say about the transmission of monetary policy? American Economic Review 90, 407–428.
Keeton, W.R., 1995. Multi-office bank lending to small businesses: some new evidence. Federal Reserve Bank of Kansas City Economic Review 81, 63–75.
Klein, B., Crawford, R.G., Alchian, A.A., 1978. Vertical integration, appropriable rents, and the competitive contracting process. Journal of Law and Economics 21, 297–326.
LaPorta, R., Lopez-de Silanes, F., Zamarripa, G., 2003. Related lending. Quarterly Journal of Economics 118, 231–268.
Mullainathan, S., Scharfstein, D., 2001. Do firm boundaries matter? American Economic Review Papers and Proceedings 91, 195–199.
Ongena, S., Smith, D.C., 2001. The duration of banking relationships. Journal of Financial Economics 61, 449–475.
Peek, J., Rosengren, E.S., 1998. Bank consolidation and small business lending: it's not just bank size that matters. Journal of Banking and Finance 22, 799–819.
Petersen, M.A., Rajan, R.G., 1994. The benefits of firm-creditor relationships: evidence from small-business data. Journal of Finance 49, 3–37.
Petersen, M.A., Rajan, R.G., 1995. The effect of credit market competition on lending relationships. Quarterly Journal of Economics 110, 407–443.
Petersen, M.A., Rajan, R.G., 2002. Does distance still matter? The information revolution in small business lending. Journal of Finance 57, 2533–2570.
Rajan, R.G., 1992. Insiders and outsiders: the choice between informed and arm's-length debt. Journal of Finance 47, 1367–1399.
Rajan, R.G., Zingales, L., 1998. Power in a theory of the firm. Quarterly Journal of Economics 113, 387–432.
Rajan, R.G., Zingales, L., 2001. The firm as a dedicated hierarchy: a theory of the origins and growth of firms. Quarterly Journal of Economics 116, 805–851.
Sapienza, P., 2002. The effects of banking mergers on loan contracts. Journal of Finance 57, 329–367.
Sharpe, S.A., 1990. Asymmetric information, bank lending and implicit contracts: a stylized model of customer relationships. Journal of Finance 45, 1069–1087.
Simester, D., Wernerfelt, B., 2003. Determinants of asset ownership: a study of the carpentry trade. Review of Economics and Statistics, forthcoming.
Stein, J.C., 2002. Information production and capital allocation: decentralized vs. hierarchical firms. Journal of Finance 57, 1891–1921.
Strahan, P.E., Weston, J.P., 1996. Small business lending and bank consolidation: is there cause for concern? Federal Reserve Bank of New York Current Issues in Economics and Finance 2, 1–6.
Strahan, P.E., Weston, J.P., 1998. Small business lending and the changing structure of the banking industry. Journal of Banking and Finance 22, 821–845.

Whinston, M.D., 2001. Assessing the property rights and transaction-cost theories of firm scope. American Economic Review Papers and Proceedings 91, 184–188.
Williamson, O.E., 1975. Markets and Hierarchies: Analysis and Antitrust Implications. Collier Macmillan, New York.
Williamson, O.E., 1985. The Economics Institutions of Capitalism. Free Press, New York.

Available online at www.sciencedirect.com
ScienceDirect

J. Finan. Intermediation 16 (2007) 452–476

Journal of Financial Intermediation

www.elsevier.com/locate/jfi

Evidence of jointness in the terms of relationship lending

Ivan E. Brick, Darius Palia [a,*]

[a] *Rutgers Business School, at Newark and New Brunswick, USA*

Received 1 December 2004

Available online 11 January 2007

Abstract

This paper examines the impact of the borrower–lender relationship on the explicit loan interest rate and collateral, as well as the correlation between loan interest rates and collateral. Using a simultaneous equation approach, we find that collateral has a statistically significant positive impact of 200 to 400 basis points on loan interest rates. We find this positive association to be stronger for personal (or outside) collateral than collateral provided by the firm's assets (or inside collateral). Finally, we find the economic impact of the borrower–lender relationship to be 21 basis points for one standard deviation increase in relationship length.

JEL classification: G20; G21

Keywords: Loan pricing; Collateral; Bank–customer relationship; Implicit interest rates

0. Introduction

The theoretical financial intermediation literature (Hodgman, 1963; Kane and Malkiel, 1965; Diamond, 1984, 1991; Ramakrishnan and Thakor, 1984; Fama, 1985; Sharpe, 1990; Rajan, 1992; and Padilla and Pagano, 1997; among others) has often argued that banks produce information about borrower firms through their lending relationships, which is otherwise not available to the capital markets. This literature emphasizes the informational advantage reaped by banks generated by repeat-business and the impact of the borrower–lender relationship upon borrowing costs.

* Corresponding author.
E-mail address: ibrick@andromeda.rutgers.edu (I.E. Brick).

1042-9573/$ – see front matter
doi:10.1016/j.jfi.2007.01.001

Researchers have, with mixed success, tested the narrower hypothesis that the length of a customer's relationship ("relationship length") significantly reduces the loan interest rate charged on its loans.[1] Empirical studies find mixed evidence for this hypothesis. Using US data, Berger and Udell (1995) find a negative correlation between relationship length and explicit interest rates whereas Petersen and Rajan (1994) find no statistical correlation.[2] The evidence using international data is also mixed. For example, Degryse and Cayseele (2000) find a positive correlation between relationship length and explicit interest rates whereas Harhoff and Korting (1998) find a negative correlation. These results may be reconciled by noting that the theory of borrower–lender relationship does not require that relationship length affect the explicit loan rate only or at all. It suggests instead that relationship value affects the sum of the implicit and explicit (or contractual) interest rates on the loan.[3] *Implicit* interest rates include, for example, additional charges or restrictions imposed on borrowers such as collateral, fees, covenanted duties, and documentation and reporting-frequency requirements. This paper examines the simultaneous impact of borrower–lender relationship upon the explicit loan interest rate and one such implicit component, namely, collateral. Since the existing theoretical and empirical literature[4] has strongly established that the collateral requirement is endogenously determined, we use a simultaneous equation approach for a sample of small business firms that are involved in relationship loans. We find that the length of the relationship does impact upon the probability of posting collateral and the level of loan interest rates. Specifically, we find the economic impact of the borrower–lender relationship to be 21 basis points for one standard deviation increase in relationship length.

Another purpose of the paper is to examine the impact of the collateral upon the level of loan interest rates. The existing literature has regressed loan interest rates on collateral with mixed results. More specifically, Berger and Udell (1995) find no statistical relationship, while Berger and Udell (1990) and John et al. (2003) find a positive relationship. These papers assume that collateral is exogenous. Our simultaneous equation approach endogenizes for collateral using valid instrumental variables that are based upon the existing literature. Using the Hansen–Sargan test (see Gujarati, 2003) we establish that our instrument variables are valid and that our system of equations is well-identified. Consistent with the empirical results of Berger and Udell (1990,

[1] Implicit in their argument is the notion that the lender learns favorable information about the borrower which results in a lower loan rate. Otherwise, the borrower could then switch to another lender that does not have the private unfavorable information.

[2] Other papers that have examined relationships in banking are: Angelini et al. (1998), Bebczuk (2004), Berger et al. (2003), Blackwell and Winters (1997), Bodenhorn (2003), Canovas and Solano (2003), Chakraborty and Hu (2006), Cole (1998), Cole et al. (2004), Conigliani et al. (1997), Cosci and Meliciani (2002), D'Auria et al. (1999), Degryse and Ongena (2005), Dennis et al. (2000), Dietsch (2003), Elsas (2005), Elsas and Krahnen (1998, 2002), Ewert et al. (2000), Ferri and Messori (2000), Guiso (2003), Hao (2003), Harhoff and Korting (1998), Lehmann and Neuberger (2001), Lehmann et al. (2004), Machauer and Weber (1998), Mallett and Sen (2001), Menkhoff and Suwanaporn (2003), Miarka (1999), Pozzolo (2004), Repetto et al. (2002), Scott and Dunkelberg (2003), Streb et al. (2002), Uzzi (1999), Weinstein and Yafeh (1998), and Ziane (2004).

[3] Kane and Malkiel (1965) portray the allocation of relationship benefits as a bilateral bargaining problem, so that benefits should be divided between banks and customers according to their relative bargaining power. A borrower with a longer relationship offers a better risk–return trade-off to the lender and therefore such a borrower can get lower rate, all else equal. At the same time, borrowers may find loan terms from non-relationship banks to be more costly because these banks lack the flow of soft information that an extended relationship generates.

[4] See, for example, Orgler (1970), Hester (1979), Besanko and Thakor (1987a, 1987b), Chan and Kanatas (1985), Bester (1985), Berger and Udell (1990), Boot et al. (1991), Boot and Thakor (1994), Berger and Udell (1995), Rajan and Winton (1995), and John et al. (2003).

1995), among others, we find that the probability of requiring firms to post collateral is positively related to firm risk as proxied by the occurrence of prior firm delinquency.

We find that when collateral is assumed to be exogenous, there is no correlation between loan rates and the probability of posting collateral. However, once we endogenize for collateral use in a simultaneous equation system, the loan rate for firms that are required to post collateral is 200–400 basis points higher than those firms that do not have to post collateral. This result is strongly consistent with John et al. (2003) that secured public debt has a higher yield than unsecured public debt. They theoretically demonstrate that agency issues between managers and lenders can give these results due to imperfections in the ratings of the credit agencies. Given that our sample consists of small non-public firms where managerial problems due to asymmetric information between insiders and external loan creditors might be higher, it is not surprising that our yield differential of collateral loans over non-collateral loan are significantly larger in magnitude than that found by John et al. (2003).

Carey et al. (1998) and Udell and Berger (2005) differentiate between asset-based lending and relationship lending. They argue that asset-based lending is more transactions oriented using "hard" information. The firm's assets, such as accounts receivable and inventory would in these cases be used as a source for collateral (also called inside collateral). In contrast, relationship lending relies on "soft" information to determine borrower risk and involve the use of personal guarantees and personal assets pledged as collateral. In addition, one might view personal collateral (also called outside collateral) as a substitute for equity, because these personal assets could be sold and the proceeds may be then used by the firm to repay the loan. John et al. (2003) examine firm (inside) collateral only, whereas other collateral papers listed in footnote 4 examine personal (outside) collateral. We distinguish between these two types of collateral in our estimation. In doing so, we separately estimate the impact of each type of collateral on the loan rate and on each other. We find that the economic impact of the requirement of posting personal (outside) collateral is greater than posting firm (inside) collateral.

Another implicit price of a loan is the amount of fees charged by lenders. In our sample, we find that the average interest rate for all lines of credit is increased by 83 basis points due to fees. However, for loans that are charged fees (433 out of a total of 766 loans), we find an average increase in the loan interest rate of 1.481% due to fees. We check whether our results are weakened or strengthened due to the inclusion of fees and we do not find that our results are materially altered.[5]

To perform our empirical tests, we use a sample of lines of credit obtained from the 1993 *National Survey of Business Finances*. Berger and Udell (1995) justify focusing on credit lines on the ground that this excludes from the "data set most loans that are 'transaction driven' rather than 'relationship driven'" (p. 353). This is not to say that relationship length has no effect on transactional loans. We focus on lines of credit to take advantage of their relative homogeneity, while recognizing that the characteristics of these loans may vary greatly. We recognize that even with our sample of lines of credit that lenders may use asset-based technology to evaluate the viability of the loan, especially once collateral requirements are imposed. We therefore perform robustness tests by restricting our sample to increase the likelihood that relationship lending based technology becomes more important in evaluating the viability of the loan to assess the

[5] In an earlier version of this paper, we also modeled the amount of fees charged as an endogenous decision. However, we found that our potential instrumental variables were also related to loan interest rates. Therefore, we are unable to accurately identify an instrumental variable for fees, and only examine whether our results are weakened or strengthened due to the inclusion of fees. We thank Jeremy Stein for suggesting this methodology.

economic impact of the borrower–lender relationship upon both the explicit interest rate and the probability of imposing a collateral requirement.

Our empirical results show strong evidence for jointness in the terms of lending. Borrowers and lenders might prefer jointness in loan pricing for two reasons. First, lenders might insist that borrowers pay implicit interest upfront as a way to reduce the lenders' interim loss exposure. The lower are the periodic interest payments, the lower the probability that bank might have to declare default.[6] Second, it might be legally and reputationally advantageous for lending institutions to mask their efforts to shade interest rates to relationship borrowers by exchanging value in implicit ways that other customers, regulators, and litigious parties cannot easily observe.

The paper proceeds as follows. Section 1 describes the data, the simultaneous-equation model that is estimated, and the various proxy and control variables that the tests employ. Section 2 presents and analyzes the empirical findings. Finally, Section 3 summarizes our results.

1. Explanation of data sources, estimation methods, and included variables

Data on borrowers and loan terms come from the 1993 *National Survey of Business Finances* (which we subsequently refer to as the *Survey*). The *Survey* is a cooperative effort of the Board of Governors of the Federal Reserve System and the Small Business Administration. We follow Berger and Udell (1995) in restricting the sample to firms who negotiate a credit line. After deleting observations with missing data, our final sample consists of 766 firms. According to the *Survey*, the initiation dates of the lines of credit range from March 1990 to December 1994. The bulk of these loans took place in 1993 (260 sample points) and 1994 (450 sample points).[7] Two macroeconomic variables are chosen to control for differences in the economic environment at the time the line was opened and are downloaded from the Federal Reserve Bank of St. Louis website (http://research.stlouisfed.org). These variables are defined later in the paper.

1.1. Limitation of survey data

The *Survey* provides only one year of borrower characteristics. This limits our ability to determine whether lenders ask a short-relationship borrower to pay implicit interest upfront as a way to reduce their interim loss exposure, or whether riskier firms borrow at fixed rates. The 1993 *Survey* includes all lines of credit, both new lines of credit and renewals of existing lines, without allowing us to distinguish which lines of credit are loan renewals and which are new initiations.[8] This limitation seems less important today than it did 18 years ago when, extending the work of James (1987), Lummer and McConnell (1989) found that loan renewals generate positive abnormal returns for borrowing firms, but loan initiations do not. Slovin et al. (1993) clarify that either loan decision generates abnormal returns, but only for small firms. Best and Zhang (1993) and Billet et al. (1995) firmly reject the hypothesis that announcement effects differ between loan initiations and renewals.

[6] To test these possibilities would require a time series of loan transactions with their associated loan rates for a specific borrower–lender relationship. This study and the extant empirical literature have not had access to such data.

[7] In addition, 2 loans from 1990, 13 loans from 1991 and 41 loans from 1992 are recorded by the *Survey*.

[8] An early version of the paper also included lines of credit recorded in the 1998 *Survey*. The 1998 *Survey* only included new lines of credit and resulted in a small sample size of 220 firms. Importantly, even in this smaller sample, there was a lot of missing data for inside and outside collateral making us drop the 1998 *Survey* from our sample.

Finally, *Survey* data does not include the name of the lender. Among others, Kashyap et al. (1993), Kashyap and Stein (2000), and Hubbard et al. (2002) argue that characteristics of banks should have an independent impact on loan contract rates. This limitation also prevents us from examining the extent to which financial-institution mergers reduce small-business lending, as documented by Berger et al. (1998) and Sapienza (2002).

1.2. Regression model

Our research compares the single-equation and simultaneous-equation estimates of the sensitivity of loan rates to collateral and the impact of relationship length on collateral and loan interest rates. Our tests will demonstrate that the loan pricing mechanisms are endogenously chosen in our sample, violating the condition needed for OLS estimation. That is because when the endogenous regressors are correlated with the regression errors, OLS estimation can yield biased estimators. Accordingly, we need to specify a simultaneous system of equations.

We begin by regressing each borrower's contract *Loan Rate Premium* on a collection of potential regressors whose relevance is supported by prior studies. We then interpose simultaneous equations for *Firm Collateral* and *Personal Collateral*. The specifications we estimate differentiate between a vector of common control variables **X** and a vector of specific instruments **Z** for each endogenous variable. The instruments are chosen to satisfy rank and order conditions that identify the following system:

$$\begin{aligned} \mathit{Loan\ Rate\ Premium} = {} & \alpha_{\mathrm{LR}} + \beta_{\mathrm{LR}} \mathit{FirmCollateral} + \gamma_{\mathrm{LR}} \mathit{PersonalCollateral} + \Omega_{\mathrm{LR}} \mathbf{X} \\ & + \lambda_{\mathrm{LR}} \mathbf{Z}_{\mathrm{LR}} + \varepsilon_{\mathrm{LR}}, \end{aligned} \tag{1}$$

$$\begin{aligned} \mathit{Firm\ Collateral} = {} & \alpha_{\mathrm{FC}} + \beta_{\mathrm{FC}} \mathit{PersonalCollateral} + \gamma_{\mathrm{FC}} \mathit{LoanRatePremium} + \Omega_{\mathrm{FC}} \mathbf{X} \\ & + \lambda_{\mathrm{FC}} \mathbf{Z}_{\mathrm{FC}} + \varepsilon_{\mathrm{FC}}, \end{aligned} \tag{2}$$

$$\begin{aligned} \mathit{Personal\ Collateral} = {} & \alpha_{\mathrm{PC}} + \beta_{\mathrm{PC}} \mathit{FirmCollateral} + \gamma_{\mathrm{PC}} \mathit{LoanRatePremium} + \Omega_{\mathrm{PC}} \mathbf{X} \\ & + \lambda_{\mathrm{PC}} \mathbf{Z}_{\mathrm{PC}} + \varepsilon_{\mathrm{PC}}. \end{aligned} \tag{3}$$

1.3. Definitions and estimators used for endogenous variables

Financial theory suggests that the required return lenders demand from borrowers should be positively related to the loan's default risk. However, borrowers can compensate lenders for risk by various combinations of loan rates and lending fees.[9] Accordingly, we use two definitions of *Loan Rate Premium*, one with fees and one without. Like in Berger and Udell (1995), *Loan Rate Premium* without fees is the difference between the contractual coupon interest set for the credit line and the prime rate for the month when the credit line is approved. *Loan Rate Premium* with fees is the difference between the effective interest rate adjusted for fees less the prime

[9] Loan fees and contract rates might relate negatively because fees can better compensate lenders for repricing options. Brueckner (1994), Dunn and McConnell (1981), Dunn and Spatt (1985), and LeRoy (1996) demonstrate theoretically that points can be used by banks to differentiate those mortgage borrowers with high probabilities of prepayment because they are more mobile from those borrowers with low probabilities of prepayment. Borrowers who are unlikely to prepay would, in equilibrium, select contracts with low interest rates and high loan fees. Borrowers who tend to prepay would select contracts with low loan fees and high interest rates. These papers envision a trade-off that is unrelated to the default risk of the firm.

rate for the month when the credit line is approved. The effective interest rate adjusted for fees is equal to the coupon rate and the ratio of total fees collected by the lending institution to the contractual amount that can be borrowed divided by the maturity denominated in months. *Firm Collateral* is a binary dummy variable that equals one if the firm is required to use its assets as collateral and is zero otherwise. *Personal Collateral* is a binary dummy variable that equals one if the CEO/owner is required to use her personal assets/guarantee as collateral and is zero otherwise. Table 1 lists these measures, along with the control and instrumental variables we employ.

Table 1
Variable definitions

Variables	Definitions
Dependent variables	
Loan Rate Premium (without fees)	the difference between the contractual coupon rate on line of credit and prime rate
Loan Rate Premium (with fees)	the difference between the effective interest rate adjusted for fees (amortized over the life of the loan) and prime rate
Firm Collateral	set to unity if the firm is required to post firm assets as collateral, and zero otherwise
Personal Collateral	set to unity if the CEO/owner is required to post personal assets/guarantee as collateral, and zero otherwise
Control variables	
Debt	ratio of total debt outstanding to level of sales
Profit	ratio of earnings before interest and taxes to sales
Cash	ratio of level of cash holdings to sales
Size	natural logarithm of firm's sales
Company	set to unity if the firm's owners enjoy limited liability protection, and zero otherwise
Term	yield spread between the five-year Treasury note and the three-month Treasury bill
Default	difference between Baa and Aaa bond yields
Relation	number of years the firm has had a relationship with the lending institution
Comp bal	set to unity if the line of credit agreement requires a compensating balance, and zero otherwise
Number	number of lending sources available to the firm
Maturity	number of months the line of credit is outstanding
Fixed	set to unity if the line of credit has a fixed coupon rate, and zero if the coupon rate is a variable rate
Instrumental variables	
For loan rate premium:	
HHI	set to unity if the Herfindahl Index for deposits in the MSA of the firm is greater than 1800, and zero otherwise
Firm age	age of the firm
Inter1	interaction of HHI and Firm age (*HHI* × *Firm age*)
For firm collateral:	
Fdelinq	set to unity if firm has previously defaulted, and zero otherwise.
For personal collateral:	
Pdelinq	set to unity if the principal owner has defaulted, and zero otherwise
CEO age	age of CEO
CEO share	percentage ownership of firm held by CEO
CEO exp	number of years experience as CEO of the firm

Control variables are chosen to meld the explanatory variables used by Berger and Udell (1995) with those used by Petersen and Rajan (1994, 1995).[10] While the *Loan Rate* regression can be estimated by the conventional least-squares techniques, the other two variables require the use of limited dependent-variable methods. As a binary variable, the range of *Firm* and *Personal Collateral* is restricted to only two values and is therefore estimated by the logistic regression.

1.4. Definition of control variables

Control variables cover four broad dimensions of the negotiation process: borrower characteristics, loan characteristics, and macroeconomic variables. Borrower characteristics are measures of the profitability and risk exposure of the firm:

(a) *Debt* (i.e., leverage), defined as the ratio of outstanding debt to annual sales;
(b) *Profit*, defined as the ratio of earnings before interest and taxes to annual sales;
(c) *Cash* (i.e., liquidity), defined as the ratio of cash holdings to the annual sales of the firm;
(d) *Size*, defined as the logarithm of annual sales;
(e) *Company*, which equals unity if the firm's owners enjoy limited-liability protection and equals zero if the firm is a sole proprietorship or partnership[11];
(f) *Number*, which proxies borrower bargaining power by the number of lending sources available to the firm at the contracting date;
(g) *Relation*, which measures the number of years the firm has dealt with the lending institution; and
(h) industry dummies which represent industry characteristics by the borrower's two-digit SIC code.[12]

Loan Rate Premium should increase with characteristics that increase the bank's exposure to loss and decrease with characteristics that increase the customer's bargaining power. On these grounds, we expect the *Loan Rate* equation to show positive coefficients for *Debt* and *Company* and negative coefficients for *Profit, Cash, Size, Relation*, and *Number*.

Three loan characteristics are also examined.

(a) *Comp Bal* takes on the value of one when the credit line must be supported by a compensating balance and is zero otherwise;
(b) *Maturity* expresses how long the credit line remains open;
(c) *Fixed* is a binary indicator that equals unity if the line of credit has a fixed coupon rate and is zero otherwise.

Macroeconomic variables represent financial conditions in the month and year the deal was completed. Two financial variables are employed:

---

[10] Although Berger and Udell (1995) test a broad range of potential determinants for borrowers' loan rates, they exclude two macroeconomic variables that Petersen and Rajan (1994, 1995) find to be significant: interest-rate variables and bank concentration ratios. Petersen and Rajan examine a wider variety of loan arrangements than Berger and Udell or ourselves, but their models suppress the role of collateral and fees.

[11] For example, *Company* is set to unity for firms that are classified as an S-Corporation, C-Corporation, Limited Liability Company or Limited Liability Partnership.

[12] To conserve space, we do not report coefficients for the industry controls.

(a) *Term*, the yield spread between the five-year Treasury note and the yield of a three-month Treasury bill;
(b) *Default*, the spread between the yields on Baa and Aaa bonds.

We expect cyclical forces to affect both the risk-free rate and market-wide risk premiums. Consequently, we expect the *Loan Rate Premium* to increase with both variables.

Although all borrower, lender, and loan characteristics are potentially endogenous, the number of endogenous variables one can handle is constrained by the difficulties of identifying a larger system. The case for endogenizing maturity and size is particularly strong (see, for example, Barclay et al., 2003; Berger et al., 2003; Brick and Ravid 1985, 1991; and Dennis et al., 2000). However, it is difficult to conceive of instruments for loan maturity and size that would not also be related to *Loan Rate Premium* and *Collateral* variables. As a robustness test, we show that our estimates of the impact of other variables do not change significantly when *Maturity* is moved into and out of each equation.[13]

1.5. Instrumental variables

Identification is achieved by drawing on prior knowledge about particular model coefficients or about the covariance matrix of the error terms. Ideally, the instrumental variables inserted into any equation should be correlated with that equation's endogenous variable, but uncorrelated with the error term (Bowden and Turkington, 1984). To compare our results with the previous literature on how collateral affects the loan rate premium, our overriding concern is to identify and obtain consistent estimates of parameters of the *Loan Rate Premium* equation (1). This leads us to introduce a large number of candidate instruments (see Chao and Swanson, 2003). Gujarati (2003) suggests that a Hansen–Sargan test for overidentifying restrictions may be used to test the exogeneity of the overall set of instruments. The null hypothesis is that the excluded exogenous variables, the instruments, are uncorrelated to the regression error. If we fail to reject the null, then we can have some confidence in the overall set of instruments used. The test is performed as follows. First we obtain the two-stage least squares residuals by regressing the independent variable against the set of control variables and the fitted values of the other independent variables. Then we regress these residuals on all the control variables and instruments and obtain its R-square. We compute the Hansen–Sargan statistic as R-square times the number observations less the number the number of regressors. The Hansen–Sargan statistic is chi-square distributed with degrees of freedom equal to the difference between the number of instruments and the number of endogenous variables. If the chi-square statistic is not significant then the instrumental variables are valid and our system of equations is well-identified.

The instrumental variables inserted into the *Loan Rate Premium* equation are measures of market power and the age of the borrowing firm. The importance of these variables is emphasized by Petersen and Rajan (1995). They argue that bankers in concentrated markets use their explicit loan rate as a loss leader to secure long-term rents on other relationship business. Petersen and Rajan believe that this strategy of loss-leadership is directed particularly at young firms. In estimating our models, *Firm age* is measured in years and we proxy market power by

[13] A similar robustness test is undertaken when *Size* is moved into and out of each equation. For brevity, results for *size* are available from the authors but not reported.

a zero-one indicator variable, *HHI*. *HHI* equals one if the Herfindahl index for deposits in the borrowers Standard Metropolitan Statistical Area exceeds 1800.[14] If *Loan Rate Premium* functions more strongly as a loss leader in concentrated markets, *HHI* should receive a negative sign. If a firm's bargaining power increases with *Firm age*, relationship rents would decline with *Firm age*, so that it should receive a negative sign. However, if loss leadership is directed especially at young firms, the interacted variable, *Inter1*—defined as the product of *Firm age* and *HHI*—should receive a positive sign.

The literature has also differentiated between personal and firm collateral. *Firm Collateral* rearranges the relative priority of liabilities upon bankruptcy without altering the risk exposure to owners. *Firm Collateral* may actually benefit the firm's owners. John et al. (2003) demonstrate that *Firm Collateral* may actually reduce expected bankruptcy costs resulting in a positive relationship between the use of *Firm Collateral*, firm risk and interest rates. We assume that the likelihood of a lender imposing firm asset collateral requirements would be positively related to whether the firm has defaulted on a previous loan. Hence, the instrumental variable that we use for *Firm Collateral* is *Fdelinq*, a dummy variable that is equal to one if the firm has previously defaulted, and zero otherwise.

Firms that are required to post firm assets as collateral might be engaged in asset-based borrowing as opposed to relationship based lending. The information technology used to evaluate asset-based loans is different from relationship based lending and therefore testing the impact of borrower–lender relationship upon loan interest rate and the probability of imposing collateral may be compromised. Accordingly, we repeat our analysis by dropping from the sample all loans that are required to post firm assets as collateral. In this way, we hope to analyze the impact of the length of the borrower–lender relationship upon a sample of loans that are evaluated primarily using relationship lending technology.

On the other hand *Personal Collateral* is a form of owner's equity and increases risk exposure to owners. The previous literature on collateral assigns it two conceptually distinct functions. On the one hand, collateral may serve to mitigate asymmetric information about borrower quality. In this case, collateral serves as a positive signal about borrower quality that is known to firm managers, but is unobservable to lenders (see, e.g., Besanko and Thakor, 1987a, 1987b; Chan and Kanatas, 1985; and Bester, 1985). On the other hand, collateral can also mitigate against moral hazard, in that it serves as a performance bond against post-loan managerial shirking and risk-taking activities that might shift uncompensated risk to the lender (see, e.g., Boot et al., 1991, and Boot and Thakor, 1994).

The above theoretical models suggest that *Personal Collateral* is used for signaling about borrower quality. To obtain a separating signaling equilibrium, the owner of the lower quality firm cannot afford to mimic the owner of higher quality firm. That is, the signal must be costly to the lower quality borrower. Similarly, *Personal Collateral* may be used to serve as a performance bond against post-loan managerial shirking and risk-taking activities because of the associated costs of managerial shirking and risk taking activities to the owner.

To account for the potential use of *Personal Collateral* as a signal for firm quality and/or as a performance bond, we introduce four instrumental variables into the *Personal Collateral* equation (3). Three of these variables are:

[14] Given that the *Survey* does not identify the lender, we cannot calculate independently the lender's Herfindahl Index. The *Survey* only provides this binary categorization.

(a) *CEO age*;
(b) *CEO share* of the ownership of the firm; and
(c) *CEO exp*, measuring the number of years the CEO has headed the firm.

We suppose that the bonding element loses importance to the lender for older and more-experienced CEOs. In addition, Bitler et al. (2004) examine the association of personal collateral and entrepreneurial wealth in testing agency theory using 2SLS and find that agency costs explain why entrepreneurs concentrate a large fraction of their wealth in their firm's equity. This implies that personal collateral is better than firm collateral to reduce agency costs since the vast majority of CEOs in our sample have less than 100% ownership. Consequently, we add a fourth instrumental variable, *Pdelinq*, a zero-one dummy that equals unity if either the principal owner or the firm has previously defaulted. We assume managerial agency gains importance if its principal owner has a history of default.[15]

Table 2 presents summary statistics for the variables included in this study. Borrower firms paid a mean and median explicit interest rate of 8.00 percent. Loan fees increased the effective rate (*erate*) by a mean of 62 basis points. For the total sample, the average fees is 83 basis

Table 2
Descriptive statistics

Variable	Unit	#	Mean	Median	St. Dev.
Dependent variables					
Loan Rate	percent	766	8.001	8.000	1.508
Erate	percent	766	8.618	8.200	2.244
Prime	percent	766	6.541	6.060	0.710
Loan Rate Premium (with fees)	percent	766	2.077	1.690	2.191
Fees (all L/Cs)	percent	766	0.831	0.065	1.963
Fees (L/Cs where fees charged)	percent	430	1.481	0.724	2.430
Firm Collateral	dummy	766	0.572	1.000	0.495
Personal Collateral	dummy	766	0.661	1.000	0.474
Control variables					
Debt	fraction	766	0.347	0.204	0.601
Profit	fraction	766	0.035	0.025	0.234
Cash	fraction	766	0.083	0.019	0.696
Size	log(Sales)	766	15.211	15.425	1.623
Sales	million	766	11.168	5.000	18.474
Company	dummy	766	0.884	1.000	0.321
Term	basis pts.	766	2.820	2.830	0.328
Default	basis pts.	766	0.700	0.680	0.077
Relation	years	766	8.535	6.000	8.412
Bank	dummy	766	0.952	1.000	0.215
Comp bal	dummy	766	0.098	0.000	0.297
Number	quantity	766	1.210	1.000	0.591
Maturity	months	766	23.548	12.000	29.124
Fixed	dummy	766	0.206	0.000	0.405

(*continued on next page*)

[15] We confirmed that our instrumental variable for *Firm Collateral* (*Fdelinq*) is not statistically related to *Personal Collateral* and our instrumental variables for *Personal Collateral* (*CEO age, CEO share, CEO exp*, and *Pdelinq*) are not statistically related to *Firm Collateral*.

Table 2 (*continued*)

Variable	Unit	#	Mean	Median	St. Dev.
Instrumental variables					
For loan rate:					
HHI	dummy	766	0.495	0.000	0.500
Firm age	years	766	18.341	14.000	16.385
Inter1	interaction	766	9.082	0.000	14.701
For firm collateral:					
Fdelinq	dummy	766	0.184	0.000	0.388
For personal collateral:					
Pdelinq	dummy	766	0.587	0.000	0.235
CEO age	years	766	51.215	50.500	10.331
CEO share	percent	766	60.424	51.000	30.296
CEO exp	years	766	21.509	20.000	10.402

Notes. This table provides summary descriptive statistics of the variables employed in our study for a sample of 766 loans (lines of credits or L/Cs) as reported by 1993 *National Survey of Business Finances*. The *Loan Rate* is the contractual coupon rate on the line of credit. *Erate* is the effective rate of the loan equaling the sum of the coupon rate and the monthly straight-line amortization rate of the fees collected by the lending institution. *Prime* is the prime rate prevailing at the initial period of the loan. *Loan Rate Premium* with fees is the difference between *Erate* and *Prime*. *Fees* is the ratio of the total fees collected by the lending institution to the total amount borrowed. The other variables are defined as follows: *Firm Collateral*, set to unity if the loan requires the firm to post its assets as collateral, and zero otherwise; *Personal Collateral*, set to unity if the owner is required to post personal assets as collateral, and zero otherwise; *Debt*, ratio of total debt outstanding to level of sales; *Profit*, ratio of earnings before interest and taxes to sales; *Cash*, ratio of the level of cash holdings to sales; *Size*, natural logarithm of firm's sales; *Sales* is the annual level of sales in millions of dollars in the year of the loan initiation; *Company*, set to unity if the firm's owners enjoy limited liability protection, and zero otherwise; *Term*, yield spread between the five-year Treasury note and the monthly three-month Treasury bill; *Default*, difference between Baa and Aaa bond yields; *Relation*, number of years the firm has had a relationship with the lending institution; *Bank*, set to unity if the lender is a bank, savings or thrift institution, and zero otherwise; *Comp bal*, set to unity if the line of credit requires a compensating balance, and zero otherwise; *Number*, number of lending institutions available to the firm; *Maturity*, number of months the line of credit is outstanding; *Fixed*, set to unity if the line of credit has a fixed coupon rate, and zero if the coupon rate is a variable rate; *HHI*, set to unity if the Herfindahl Index for deposits in the MSA of the firm is greater than 1800, and zero otherwise; *Firm age*, age of the firm; and *Inter1*, interaction of HHI and Firm age (HHI × *Firm age*); *Fdelinq*, set to unity if the firms has previously defaulted, and zero otherwise; *CEO age*, age of the CEO; *CEO share*, percentage ownership of firm held by CEO; *CEO exp*, number of years experience as CEO of the firm. *Pdelinq*, set to unity if the principal owner has defaulted, and zero otherwise.

points, but the median effect is only 6.5 basis points. This is because only 430 of the 766 sample borrowers paid fees at all. For the fee-paying firms, the net contribution by fees averaged 148 basis points, with a median of 72.4 basis points. Fully 57.2 percent and 66.1 percent of sample lines of credit included, respectively, firm collateral personal collateral requirements. Borrower sales averaged $11.17 million, while median sales were only $5.0 million. Debt averaged 34.7 percent of sales. Profits averaged 3.5 percent of sales, but the median ratio was only 2.5 percent. Firms held an average of 8.30 percent of their sales in cash.

Across sample dates, the mean prime rate was 6.54 percent, while the median rate was a slightly lower 6.06 percent. The mean monthly-average yield spread (*Term*) between the five-year Treasury notes and Treasury bills averaged 2.82 percent. The yield difference between the Baa and Aaa bonds (interpreted as a market-wide default premium) averaged 0.70 percent. Relationship length (*Relation*) showed a mean of 8.54 years and a median of six years. The mean age of sample firms is 18.34 years.[16] On average, 95.2 percent of the credit lines were written by deposit institutions and 9.8 percent of lines carried a compensating balance. The mean maturity of sample lines was 23.55 months, with a median maturity of only one year.

Fully 49.5 percent of the firms borrowed from institutions in highly concentrated markets implying that the median firm borrowed in a competitive market. Nearly 20.6 percent of the lines carried a fixed interest rate, which implies that the median firm accepted a variable-rate contract. Firms averaged only 1.21 sources for lines of credit and the median firm showed only one source. The mean CEO is 51 years old, with 21.5 years of experience. The mean (median) share of equity owned by the CEO was 60.42 percent (51 percent). Approximately 18.40 percent of sample firms showed a record of prior default. In contrast, almost 60% of the owners have previously defaulted on a personal obligation.

2. Empirical evidence of the interaction of implicit and explicit loan terms

To establish a benchmark with which to evaluate simultaneous-equation results, Table 3 reports single-equation estimates for two models of the *Loan Rate Premium*. The models differ only in whether or not the loan rate premium is adjusted for fees charged. We also performed the regressions with and without the inclusion of *Maturity* as an independent variable. Although not formally reported, we find that the estimates of the other coefficients and overall model performance are not generally affected by whether or not we include *Maturity*.

Neither single-equation specification provides convincing evidence that loan rates and collateral are correlated. In Table 3, only *Personal Collateral* is even marginally significant (and only when the loan rate premium is not adjusted for fees). Because the dummy variable *Personal Collateral* (and *Firm Collateral*) equals one only if there exists a collateral requirement, the 0.199 coefficient for *Personal Collateral* implies that on loans that require personal collateral the premium runs approximately 20 basis points higher than loans that do not require personal collateral. That is, this single-equation model generates no strong evidence that explicit interest rates and implicit interest rates are correlated, and accordingly no indication that they are jointly determined.

Note that *Profit, Size, Company, Relation, Fixed, Pdelinq* and *Inter1* prove statistically significant when we adjust the *Loan Rate Premium* for fees. *Debt, Profit, Fixed* and *Pdelinq* prove significant at five-percent, while *Debt, Comp bal, Fdelinq, HHI* and *Inter1* are significant at weaker levels. These estimates indicate, as in the previous literature, that smaller borrower firms are charged a higher contract rate. Interestingly, borrowing firms with limited-liability enjoy a 63 basis point savings on loans compared to borrowing firms that do not have limited liability protection. Our results also show that owners of firms who previously defaulted pay an economically significant premium differential of approximately 100 basis points. Unlike Petersen and Rajan (1994, 1995), contract rates do not vary significantly with macroeconomic variables. Our estimates are consistent with those of Berger and Udell (1995) in finding a significant negative association between *Loan Rate Premium* adjusted for fees and the length of the relationship the borrower has had with the lending institution. However, the coefficient is not economically significant (2.5 basis point reduction per year, or 21 basis points reduction for one standard deviation

[16] Note that this average age is somewhat higher than the average age of 14.1 years in the Berger and Udell (1995) sample. However, the average length of relationship in our sample is only 8.54 years, which is well below the Berger–Udell mean of 11.4 years.

Table 3
Single-equation regression of the Loan Rate Premium (all regressors treated as exogenous)

Variable	With fees		Without fees	
	parameter	(*p*-value)	parameter	(*p*-value)
Intercept	6.297[a]	(0.001)	4.172[a]	(0.000)
Firm Collateral	0.125	(0.438)	0.076	(0.463)
Personal Collateral	0.217	(0.191)	0.199[c]	(0.063)
Debt	−0.089	(0.550)	−0.173[c]	(0.077)
Profit	−0.687[c]	(0.051)	−0.479[b]	(0.035)
Cash	−0.042	(0.734)	−0.082[b]	(0.035)
Size	−0.229[a]	(0.000)	−0.149[a]	(0.000)
Company	−0.631[b]	(0.017)	−0.214	(0.210)
Term	0.033	(0.888)	−0.038	(0.798)
Default	0.660	(0.507)	−0.282	(0.662)
Relation	−0.025[b]	(0.017)	−0.007	(0.285)
Comp bal	0.023	(0.930)	−0.279[c]	(0.094)
Number	−0.030	(0.821)	−0.033	(0.698)
Maturity	−0.004	(0.124)	0.002	(0.196)
Fixed	0.498[a]	(0.001)	0.416[a]	(0.001)
Fdelinq	0.271	(0.199)	0.262[c]	(0.055)
Pdelinq	1.134[a]	(0.001)	0.933[a]	(0.000)
HHI	−0.351	(0.123)	−0.264[c]	(0.073)
Firm age	−0.001	(0.337)	−0.003	(0.484)
Inter1	0.019[b]	(0.034)	0.011[c]	(0.064)
CEO share	0.004[c]	(0.089)	0.003[c]	(0.097)
CEO age	−0.016	(0.127)	0.003	(0.806)
CEO exp	−0.002	(0.881)	−0.005	(0.421)
Adj. R^2	0.111		0.115	

Notes. This table provides the ordinary-least-squares regression results of *Loan Rate Premium* on all exogenous variables specified in Table 2. The *Loan Rate Premium* without fees is defined as the difference between contractual coupon on the line of credit and the prime rate. The *Loan Rate Premium* with fees increases the coupon rate by adding the monthly straight-line amortization rate of the fees collected by the lending institution as a fraction of the total amount borrowed over the life of the loan. The independent variables are defined as in Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.

[a] Statistically significant at 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

increase in relationship length) and is no longer significant when the *Loan Rate Premium* without adjusting for fees is the dependent variable. This last result is consistent with the results of Petersen and Rajan (1994) who find no significant association between relationship length and explicit interest rates.

We next investigate the effects of treating *Personal* and *Firm Collateral* as determined simultaneously with the *Loan Rate Premium*. We begin by assessing the strength of our instrumental variables for our collateral variables in Tables 4 and 5. Tables 4 and 5 report results for two specifications. The first specification measures the explanatory power of the instrumental variables alone. The second specification incorporates the contribution of the instrumental variables and control variables taken together. We assess the strength of our instrumental variables for the *Loan Rate Premium* in Table 6. In this table, we have four specifications. The first specification measures the explanatory power of the instrumental variables alone while the second incorporates the contribution of the instrumental variables and control variables taken together when the

Loan Rate Premium adjusted for fees is the dependent variable. The next two specifications are analogous to the first two, except now the dependent variable is the *Loan Rate Premium* with no adjustment for fees.

Table 4 summarizes the logistic regression for *Firm Collateral*. Depending on its sign, each significant coefficient increases or decreases the probability that the borrower is required to post collateral. The first specification establishes a significantly positive relationship between *Firm Collateral* and *Fdelinq*. The experiment achieves a significant likelihood ratio of 40.21. In the table's second specification, the coefficient of *Fdelinq* remains significant. Many of the control variables are significant. In particular, we find that the probability the lender will impose a requirement that the firm posts assets for collateral increases with firm leverage, firm size and whether the borrower is incorporated. The probability of firm collateral requirement decreases with the length of relationship between the borrower and the lending institution. However, the economic impact is not larges as the odds ratio is close to one, implying that a one-year increase in borrower–lender relationship results in an almost equal probability of posting collateral and not posting collateral.

In parallel fashion, Table 5 assesses the performance of the instruments included in the *Personal Collateral* equation. In the first specification, the instruments *CEO age* and *CEO share* prove significant. When we include control variables in the second specification, the results on *CEO age* and *CEO share* remain statistically significant and a χ^2 test that the instruments are jointly equal to zero is rejected. Among the controls, *Size, Company, Term* and *Fixed* prove significant. In particular, the likelihood of a personal collateral requirement decreases with the

Table 4
Evidence of the validity of instruments for Firm (Inside) Collateral

Variable	Instrumental variables only			Instrumental and control variables		
	parameter	(p-value)	odds ratio	parameter	(p-value)	odds ratio
Intercept	−0.139	(0.382)		−1.797	(0.164)	
Instrumental variable:						
Fdelinq	0.449[b]	(0.026)	1.566	0.470[b]	(0.025)	1.600
Control variables:						
Debt				0.398[b]	(0.015)	1.489
Profit				0.249	(0.492)	1.283
Cash				−0.012	(0.926)	0.989
Size				0.152[a]	(0.005)	1.165
Company				0.731[a]	(0.006)	2.078
Term				−0.181	(0.442)	0.834
Default				−1.151	(0.256)	0.316
Relation				−0.025[a]	(0.008)	0.975
Comp bal				−0.141	(0.590)	0.869
Number				0.143	(0.313)	1.154
Maturity				0.004	(0.152)	1.004
Fixed				−0.164	(0.395)	0.848
Likelihood ratio	40.21[a]	(0.000)		81.17[a]	(0.000)	

Notes. This table provides the logistic regression results explaining the requirement of firm's assets used as collateral to borrowers on lines of credit for a sample of 766 firms as reported by 1993 *National Survey of Business Finances*. For dependent variables, see Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.

[a] Statistically significant at 1% level.

[b] Idem, 5%.

Table 5
Evidence of validity of instruments for Personal (Outside) Collateral

Variable	Instrumental variables only			Instrumental and control variables		
	parameter	(p-value)	odds ratio	parameter	(p-value)	odds ratio
Intercept	1.503[a]	(0.002)		6.050[a]	(0.000)	
Instrumental variables						
Pdelinq	0.602	(0.110)	1.826	0.522	(0.173)	1.686
CEO age	−0.025[b]	(0.016)	0.975	−0.022[b]	(0.039)	0.978
CEO share	0.009[a]	(0.001)	1.009	0.008[a]	(0.007)	1.008
CEO exp	0.006	(0.534)	1.006	0.012	(0.250)	1.012
Control variables						
Debt				0.076	(0.683)	1.079
Profit				0.272	(0.484)	1.312
Cash				0.202	(0.633)	1.223
Size				−0.205[a]	(0.001)	0.814
Company				0.631[b]	(0.025)	1.879
Term				−0.540[b]	(0.032)	0.582
Default				−0.902	(0.387)	0.406
Relation				−0.012	(0.222)	0.988
Comp bal				0.157	(0.568)	1.170
Number				0.160	(0.276)	1.173
Maturity				−0.001	(0.606)	0.999
Fixed				−0.452[b]	(0.026)	0.636
χ^2-value that instruments are jointly = 0				2.172[c]	(0.099)	
Likelihood ratio	44.301[a]	(0.000)		71.704[a]	(0.000)	

Notes. This table provides the logistic regression results explaining the requirement of owner's assets used as collateral to borrowers on lines of credit for a sample of 766 firms as reported by 1993 *National Survey of Business Finances*. For dependent variables, see Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.
[a] Statistically significant at 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

size of the firm, is higher if the firm enjoys limited liability status, decreases with the spread between the five-year Treasury note and three-month Treasury Bill and is lower if the coupon rate is fixed. The logistic regression shows a significant likelihood ratio in both specifications. Finally, we do not see a statistical impact of *Relation* on the probability of posting personal collateral.

In Table 6, the instrumental variables for *Loan Rate Premium* prove strongly significant in the first and third specifications, but their significance weakens when control variables are introduced in the second and fourth specifications. Nevertheless, an F-test rejects the hypothesis that instrumental variable coefficients are jointly equal to zero. Several of the control variables prove significant in the second and fourth specifications, including *Profit, Size* and *Fixed*.

In these single-equation experiments, coefficient signs are only partly consistent with the results of Petersen and Rajan (1995). The negative coefficient for *HHI* continues to indicate that banks in more-concentrated markets may charge a lower explicit rate. However, when we include fees, the coefficients for *HHI* are negative but are not significant. In addition, the negative sign for *Firm age* implies—in line with the bargaining-power hypothesis—that older firms receive

Table 6
Evidence of validity of instruments for the Loan Rate Premium

Variable	With fees		With fees		Without fees		Without fees	
	parameter	(p-value)	parameter	(p-value)	parameter	(p-value)	parameter	(p-value)
Intercept	2.766[a]	(0.000)	7.120[a]	(0.000)	1.875[a]	(0.000)	5.186[a]	(0.000)
Instrumental variables								
HHI	−0.286	(0.228)	−0.340	(0.139)	−0.257[c]	(0.096)	−0.271[c]	(0.070)
Firm age	−0.023[a]	(0.001)	−0.010	(0.155)	−0.013[a]	(0.004)	−0.005	(0.297)
Inter1	0.014	(0.156)	0.017[c]	(0.065)	0.010	(0.107)	0.011[c]	(0.077)
Control variables								
Debt			−0.082	(0.587)			−0.149	(0.129)
Profit			−0.703[b]	(0.048)			−0.481[b]	(0.038)
Cash			−0.078	(0.537)			−0.112	(0.173)
Size			−0.287[a]	(0.000)			−0.189[a]	(0.000)
Company			−0.572	(0.031)			−0.164	(0.340)
Term			−0.008	(0.971)			−0.085	(0.577)
Default			0.470	(0.640)			−0.390	(0.552)
Relation			−0.029[a]	(0.005)			−0.010	(0.145)
Comp bal			−0.063	(0.807)			−0.336[c]	(0.048)
Number			0.013	(0.924)			−0.004	(0.961)
Maturity			−0.004	(0.125)			0.002	(0.254)
Fixed			0.512[a]	(0.008)			0.430[a]	(0.001)
F-value that instruments are jointly = 0			2.340[c]	(0.100)			3.360[c]	(0.067)
Adj. R^2	0.017		0.110		0.009		0.076	

Notes. This table summarizes the results of the ordinary-least-squares regression of the *Loan Rate Premium* variable for a sample of 766 firms as reported by 1993 *National Survey of Business Finances*. For dependent variables, see Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.

[a] Statistically significant at 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

the lowest explicit rates. Finally, the loan rate spread that includes fees is significantly negatively related to *Relation*. However, the coefficient on *Relation* implies that the loan rate decreases by only 3 basis points per additional year of *Relation*.

Table 7 reports two-stage least-squares (2SLS) estimates of the *Loan Rate* equation using the three-equation system. In the first-stage, we regress *Personal* and *Firm Collateral* (logistic regression) on all instrumental variables and control variables. Fitted values from this stage are used as regressors in *Loan Rate* specifications that alternately include and exclude fees.

Finally, we compute the Hansen–Sargan statistic to test the null hypothesis that the excluded exogenous variables, the instrument variables, are uncorrelated to the regression error. The Hansen–Sargan statistic is chi-square distributed with degrees of freedom equal to the difference between the number of instruments and the number of endogenous variables. The last row of Table 7 provides the Hansen–Sargan statistic demonstrating that the instruments are uncorrelated to the regression error, thereby formally confirming the validity of our instrumental variables and that our system of equations is accurately identified.

We observe that the endogenous variable *Personal Collateral* becomes economically significant in the 2SLS regressions. Recall that in Table 2 the interest rate on loans with a personal collateral requirement were only 20 basis points higher than that of loans without personal loan

Table 7
Simultaneous system of equation estimation for the Loan Rate Premium

Variable	With fees		Without fees	
	parameter	(p-value)	parameter	(p-value)
Intercept	−0.991	(0.707)	−0.555	(0.746)
Firm Collateral (fitted)	0.791	(0.538)	0.346	(0.679)
Personal Collateral (fitted)	4.434[a]	(0.001)	3.148[a]	(0.000)
HHI	−0.396[c]	(0.087)	−0.304[b]	(0.044)
Firm age	0.000	(0.996)	0.002	(0.707)
Inter1	0.027[a]	(0.005)	0.018[a]	(0.005)
Debt	−0.153	(0.424)	−0.180	(0.150)
Profit	−0.983[a]	(0.007)	−0.667[a]	(0.005)
Cash	−0.133	(0.294)	−0.152[c]	(0.065)
Size	−0.107	(0.289)	−0.054	(0.414)
Company	−1.188[a]	(0.001)	−0.568[b]	(0.011)
Term	0.502[c]	(0.063)	0.268	(0.126)
Default	1.779[c]	(0.099)	0.485	(0.490)
Relation	−0.024[b]	(0.037)	−0.007	(0.358)
Comp bal	−0.140	(0.596)	−0.397[b]	(0.021)
Number	−0.188	(0.189)	−0.140	(0.132)
Maturity	−0.004	(0.177)	0.002	(0.210)
Fixed	0.960[a]	(0.000)	0.740[a]	(0.000)
Hansen–Sargan χ^2-value that $E(\varepsilon_{\text{LR}}, Z) = 0$	9.485	(0.148)	11.19	(0.138)
Adj. R^2	0.099		0.092	

Notes. This table presents the results of the two-stage least-squares (2SLS) regression of *Loan Rate Premium* variables treating *Firm Collateral* and *Personal Collateral* as endogenously determined for a sample of 766 firms as reported by 1993 *National Survey of Business Finances*. For dependent variables, see Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.

[a] Statistically significant at 1% level.

[b] Idem, 5%.

[c] Idem, 10%.

collateral requirements. With 2SLS, we use the fitted values for our two collateral variables. These fitted values represents the probability that a firm will be required to post collateral and no longer is a dichotomous zero-one variable. We interpret the coefficient on the collateral variables to imply the differential loan rate premium between those loans with a 100% probability of collateral imposition and those loans with a zero percent probability of collateral imposition. We find that the coefficient for *Personal Collateral* implies a differential that is between 315 and 443 basis points. The significantly positive signs that simultaneous modeling assigns to this variable support the hypothesis that posting *Collateral* controls and implicitly prices some (but not all) of the loss exposure lenders face in high-risk loans.

In Table 7, we observe that the coefficients for *HHI* are significantly negative and the coefficients for *Inter1* are significantly positive. The positive sign for the interaction term *Inter1* supports the Petersen–Rajan hypothesis that, in concentrated markets, banks may charge younger firms a lower explicit rate on the grounds they can confidently plan to extract rents from these customers in other ways or at other times. However, when we exclude the interaction term *Inter1* in both specifications we find *HHI* to be positive but not statistically significant and *Firm age* to be positive and significant at the 10% level. None of our other results are affected and therefore we only report the results with the interaction variable.[17] The negative coefficients for *Profit* and *Company* imply that profitable and limited liability lenders are able to borrow at lower interest rates than their counterparts.

The economic importance of the length of the relationship between the lender and borrower does not become any more important with simultaneous estimation. There are several potential explanations for this. First, Berger and Udell (1995) find that the relationship variable is only significant for larger lines of credit, whereas our sample includes lines of credit as small as $10,000. Small lines of credit loans are very likely to be micro-business loans and/or personal loans that depend more on the entrepreneur's personal credit rather than the length of the relationship between the firm and the lender. Second, lines of credit of very large size and/or long maturities might finance large projects, which may use asset-based lending technology. Third, Berger and Udell (1995) point out that asset-based lending uses hard information, based upon the value of the firm collateral to assess the viability of the loan. In such a case, the length of the borrower–lender relationship becomes less important. Given that more than 60 percent of our borrowers post collateral our tests are biased against finding any significant impact of relationship length. Fourth, our sample includes loans from commercial finance companies that are primarily asset based lenders.

Consequently, to test the robustness of our results, as well as to see if *Relation* becomes more important, we modify our sample to increase the likelihood that the loans in our sample are relationship based. First, we restrict the size of the line of credit to be greater than $250,000 and less than $25 million. Second, we exclude from the sample any loan with maturity greater than one year. Third, we exclude any loan that requires the firm to post collateral. Fourth, we exclude from the sample loans obtained from finance companies. The results of these robustness tests are summarized in Table 8. Since the results we obtain are similar whether or not we adjust for fees, the table only reports the results when the loan rate premium is adjusted for fees.

The first specification of Table 8 reports the results when the sample only includes lines of credit in excess of $250,000 and less than $25 million. Note that this restriction reduces our original sample of 766 firms to 444 firms. In this case, the coefficient for *Relation* is no longer statistically significant. Further note, that the coefficients for both *Personal* and *Firm Collateral* variables are both economically and statistically significant. The second specification adds the restriction that we exclude loans with maturity greater than one year. Note that the sample is further reduced to 325 firm observations. In this case, the coefficient for *Relation* remains statistically insignificant. We observe that the coefficients for *Firm Collateral* is positive but not significant (p-value equal to 11%) whereas *Personal Collateral* remains strongly significant. The third specification keeps the size restriction and excludes all loans that require the firm to post collateral. In this case, we are only simultaneously estimating two equations, the *Loan Rate Premium* and *Personal Collateral* equations. Note that the sample is now 328 firm observations. We see that the *Personal Collateral* variable remains both economically and statistically significant. The *Relation* variable is now significant but it is not economically significant since the results imply that a borrower with a ten-year relationship with the lending institution will find that the loan will carry lower interest rate of 36 basis points compared to a newcomer. In the fourth specification, we exclude loans from finance companies without change in results that we obtained earlier. This is not surprising in that only 37 of the 766 loans of our total sample were originated by finance companies. Finally, in the fifth specification we restrict our sample to lines of credit greater than

[17] The results without the interaction variable are available from the authors upon request.

Table 8
Simultaneous system of equation estimation for the Loan Rate Premium with fees: Robustness Tests

Variable	$N = 444$ Size restriction		$N = 325$ Size restriction and maturity < 1 year		$N = 328$ Size restriction and no firm collateral		$N = 729$ No finance companies		$N = 419$ No finance companies and size restriction	
	parameter	(p-value)	parameter	(p-value)	parameter	(p-value)	parameter	(p-value)	parameter	(p-value)
Intercept	−1.419	(0.586)	−2.177	0.445	−1.222	(0.756)	2.916	(0.272)	0.058	(0.981)
Firm Collateral (fitted)	2.769[b]	(0.017)	1.434	0.107			1.362	(0.270)	1.969[b]	(0.031)
Personal Collateral (fitted)	3.726[a]	(0.000)	2.626[b]	0.013	5.219[a]	(0.012)	2.456[c]	(0.056)	2.751[a]	(0.002)
HHI	−0.654[b]	(0.018)	−0.393	0.173	−0.420	(0.292)	−0.407[c]	(0.067)	−0.465[c]	(0.052)
Firm age	0.006	(0.459)	−0.010	0.225	0.013	(0.336)	−0.004	(0.629)	−0.002	(0.826)
Inter1	0.020[b]	(0.034)	0.021[b]	0.045	0.029[c]	(0.068)	0.025[a]	(0.007)	0.019[b]	(0.023)
Debt	−0.062	(0.781)	−0.045	0.872	−0.190	(0.491)	−0.172	(0.352)	0.031	(0.871)
Profit	−1.558[a]	(0.002)	−0.939	0.126	−1.019[c]	(0.076)	−0.803[b]	(0.019)	−0.825[c]	(0.063)
Cash	0.076	(0.931)	0.155	0.878	0.085	(0.731)	−0.120	(0.314)	−0.261	(0.719)
Size	0.115	(0.274)	0.127	0.338	−0.080	(0.525)	−0.264[b]	(0.013)	0.032	(0.740)
Company	−0.993[b]	(0.037)	−1.089[b]	0.035	−1.229[a]	(0.007)	−1.143[a]	(0.001)	−0.928[b]	(0.020)
Term	0.017	(0.955)	0.253	0.489	0.728[c]	(0.066)	0.374	(0.140)	0.060	(0.817)
Default	−1.684	(0.154)	0.447	0.817	0.488	(0.789)	1.593	(0.118)	−0.874	(0.387)
Relation	−0.009	(0.433)	−0.008	0.528	−0.036[b]	(0.034)	−0.019[c]	(0.085)	−0.007	(0.472)
Comp bal	−0.018	(0.943)	−0.010	0.971	−0.158	(0.731)	−0.087	(0.735)	−0.073	(0.746)
Number	−0.463[a]	(0.009)	−0.409[b]	0.042	−0.514	(0.108)	−0.188	(0.196)	−0.364[b]	(0.023)
Maturity	0.000	(0.962)	–		−0.011[c]	(0.058)	−0.006[b]	(0.044)	−0.001	(0.754)
Fixed	0.955[a]	(0.000)	0.874[a]	0.004	1.175[a]	(0.001)	0.882[a]	(0.000)	0.937[a]	(0.000)
Adj. R^2	0.078		0.072		0.089		0.123		0.105	

Notes. This table presents the results of the two-stage least-squares (2SLS) regression of *Loan Rate Premium* variables treating *Firm Collateral* and *Personal Collateral* as endogenously determined for various sub-samples of the original sample of 766 firms as reported by 1993 *National Survey of Business Finances*. For dependent variables, see Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.

[a] Statistically significant at 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

$250,000 but less than $25 million, and loans originated by deposit institutions. Again the collateral results are similar to that of the second specification but now the *Relation* variable is no longer significant. The lack of strong results for *Relation* might be explained theoretically by the presence of strong competition among lenders as explained by Padilla and Pagano (1997) and von Thadden (2004).

In the first three columns of Table 9, we examine the impact on *Firm Collateral* when we include the other endogenous variables, *Personal Collateral* and the *Loan Rate Premium*, as regressors.[18] In the final three columns, we estimate the impact on *Personal Collateral* when we include the other endogenous variables, *Firm Collateral* and the *Loan Rate Premium*, as regressors. We find no effect of the explicit *Loan Rate Premium* on both types of collateral. This is not surprising in light of the fact that lenders might set a limit as to the maximum interest rate level that is charged. This rate might be set to reduce lenders' interim exposure because higher interest rates will increase the borrower's probability of default. Instead, high-risk borrowers

Table 9
Simultaneous system of equation estimation for Collateral with fees

Variables	Firm Collateral			Personal Collateral		
	parameter	(p-value)	odds ratio	parameter	(p-value)	odds ratio
Intercept	−3.175	0.190		8.229	(0.011)[b]	
Loan Rate Premium (fitted)	−0.128	0.607	0.880	−0.477	(0.305)	0.621
Firm Collateral (fitted)				3.118	(0.050)[b]	22.602
Personal Collateral (fitted)	1.221	0.310	3.390			
Fdelinq	0.479	0.045	1.614			
Pdelinq				1.191	(0.099)[c]	3.290
Ceo_age				−0.017	(0.189)	0.983
Ceo_share				0.011	(0.004)[a]	1.011
Ceo_exp				0.012	(0.278)	1.012
Debt	0.386	(0.020)[b]	1.471	−0.268	(0.291)	0.765
Profit	0.104	(0.802)	1.110	−0.216	(0.683)	0.806
Cash	−0.032	(0.799)	0.968	0.217	(0.634)	1.242
Size	0.181	(0.042)[b]	1.199	−0.429	(0.006)[a]	0.651
Company	0.511	(0.157)	1.666	−0.148	(0.778)	0.862
Term	−0.048	(0.857)	0.953	−0.393	(0.135)	0.675
Default	−0.835	(0.432)	0.434	0.070	(0.953)	1.072
Relation	−0.025	(0.028)[b]	0.976	−0.009	(0.516)	0.991
Comp bal	−0.177	(0.504)	0.838	0.253	(0.366)	1.288
Number	0.098	(0.508)	1.103	0.064	(0.681)	1.066
Maturity	0.004	(0.208)	1.004	−0.006	(0.129)	0.994
Fixed	0.008	(0.976)	1.008	−0.121	(0.703)	0.886
Likelihood ratio	82.2	(0.001)[a]		75.72	(0.001)[a]	

Notes. This table presents the results of the two-stage least-squares (2SLS) regression of *Firm Collateral* or *Personal Collateral* treating the other two dependent variables as endogenous for a sample of 766 firms as reported by 1993 *National Survey of Business Finances*. For dependent variables, see Table 2. p-values are given in parentheses. Two-digit industry dummies are included, but their coefficients are not reported.

[a] Statistically significant at 1% level.
[b] Idem, 5%.
[c] Idem, 10%.

[18] Table 9 reports only for the *Loan Rate Premium* adjusted for fees. The results are not affected when we do not include fees.

can adequately compensate lenders by posting collateral. Alternatively, it might be legally and reputationally advantageous for lending institutions to mask their efforts to shade interest rates to relationship borrowers by exchanging value in implicit ways that other customers, regulators, and litigious parties cannot easily observe. Consequently, at or near that maximum rate, there will be borrowers who will not need collateral to adequately compensate the lender as well as borrowers who will need to post collateral. Therefore the loan rate does not determine the posting of collateral. We also observe that the economic impact of *Relation* on collateral using simultaneous equations is not any stronger than the results obtained in single-equation framework (Tables 4 and 5).

3. Conclusions

The theoretical financial intermediation literature has often argued that banks produce information about borrower firms through their lending relationships, which is otherwise not available to the capital markets. Researchers have, with mixed success, tested the narrower hypothesis that the length of a customer's relationship ("relationship length") significantly reduces the loan interest rates. In addition the empirical literature finds mixed results with respect to the correlation between collateral requirements and loan interest rates.

These mixed results may be reconciled by noting that the theory of bank–customer relationship does not require that relationship length affect the explicit loan rate only or at all. It suggests instead that relationship value affects the sum of the explicit and implicit interest rates such as the requirement to post collateral. This paper examines the impact of borrower–lender relationship upon the explicit loan interest rate and collateral. Since the existing theoretical and empirical literature has strongly established that collateral requirement is endogenously determined, we use a simultaneous equation approach for a sample of small business firms that are involved in relationship loans. We find that the length of the relationship does impact upon both the probability of posting collateral and the level of the loan interest rates. Specifically, we find the economic impact of the borrower–lender relationship to be 21 basis points for one standard deviation increase in relationship length. We also find that collateral has a statistically significant economic impact of 200 to 400 basis points on loan interest rates, suggesting strong evidence for jointness in the terms of lending. Further, consistent with the moral hazard theories of collateral and the results of Berger and Udell (1990, 1995), we find a positive correlation between observable firm risk variables and posting collateral.

Another implicit price of a loan is the amount of fees charged by lenders. In our sample, more than half of the loans (433 out of a total of 766 loans) are charged fees. Fees should be endogenously determined along with the requirement of posting collateral. However we found that our potential instrumental variables for fees were also related to loan interest rates. Therefore, we are unable to accurately identify an instrumental variable for fees, and only examine whether our results are weakened or strengthened due to the inclusion of fees. We do not find that our results are materially altered by the inclusion or exclusion of fees in our analysis.

Carey et al. (1998) and Udell and Berger (2005) differentiate between asset-based lending and relationship lending. They argue that asset-based lending is more transactions oriented using "hard" information. The firm's assets, such as accounts receivable and inventory would in these cases be used as a source for collateral (also called inside collateral). In contrast, relationship lending relies on "soft" information to determine borrower risk and involve the use of personal guarantees and personal assets pledged as collateral. Further, one might view personal collateral (also called outside collateral) as a substitute for equity, because these personal assets could be sold and the proceeds may be then used by the firm to repay the loan. We distinguish between these two types of collateral in our estimation. In doing so, we separately estimate the impact of each type of collateral on the loan rate and on each other. We find that the economic impact of the requirement of posting personal (outside) collateral is greater than posting firm (inside) collateral.

Future research might examine the impact of lender characteristics on loan interest rates (as in Kashyap et al., 1993; Kashyap and Stein, 2000; and Hubbard et al., 2002) while taking into account self-selection among the borrowers and lenders. In addition, one might examine the role of foreign banks on loan interest rates, collateral, and fees. Maybe they prefer a different borrower profile from US domestic banks. We leave such issues for future research.

Acknowledgments

We thank Allan Berger, N. Chidambaran, William Greene, Kose John, Ed Kane, Michael Long, Michael Pagano, Bob Porter, Matt Richardson, Alan Richert, George Pennacchi, Ben Sopranzetti, Philip Strahan, Jeremy Stein, Anjan Thakor, Greg Udell, Ernst-Ludwig von Thadden, Chia-Jane Wang, Robert Whitelaw, Wei Yu, participants at the 2004 Midwest Finance Association Meetings, the 2004 Hawaii International Conference on Business, 2005 American Finance Association Meetings and seminars at Baruch, Connecticut, Hofstra, IUPUI, Maryland, NYU, Vanderbilt, and Yale for helpful comments and discussions, and the Whitcomb Center for Research in Financial Services for financial support. We especially thank two anonymous referees and George Pennacchi (Editor) whose comments have significantly improved the paper. All errors remain our responsibility.

References

Angelini, P., Di Salvo, R., Ferri, G., 1998. Availability and cost of credit for small businesses: Customer relationships and credit cooperatives. J. Banking Finance 22, 925–954.

Barclay, M., Marx, M.L., Smith, C., 2003. The joint determination of leverage and maturity. J. Corp. Finance 9, 149–167.

Bebczuk, R.N., 2004. What determines the access to credit by SMEs in Argentina? Unpublished manuscript. Universidad Nacional de la Plata.

Berger, A., Udell, G., 1990. Collateral, loan quality, and bank risk. J. Monet. Econ. 25, 21–42.

Berger, A., Udell, G., 1995. Relationship lending and lines of credit in small firm finance. J. Bus. 68, 351–381.

Berger, A., Saunders, A., Scalise, J.M., Udell, G., 1998. The effect of bank mergers on small business lending. J. Finan. Econ. 50, 187–229.

Berger, A., Rosen, R.J., Udell, G., 2003. Does market size structure affect competition? The case of small business lending. Unpublished manuscript. Board of Governors of the Federal Reserve System, Washington, DC.

Besanko, D., Thakor, A., 1987a. Collateral and rationing: Sorting equilibria in monopolistic and competitive credit markets. Int. Econ. Rev. 28, 671–689.

Besanko, D., Thakor, A., 1987b. Competitive equilibrium in the credit market under asymmetric information. J. Econ. Theory 71, 291–299.

Best, R., Zhang, H., 1993. Alternative information sources and the information content of bank loans. J. Finance 48, 1507–1522.

Bester, H., 1985. Screening vs. rationing in credit markets with imperfect information. Amer. Econ. Rev. 75, 850–855.

Billet, M.T., Flannery, M.J., Garfinkel, J.A., 1995. The effect of lender identity on a borrowing firm's equity return. J. Finance 50, 699–718.

Bitler, M., Moskowitz, T.J., Vissing-Jorgensen, A., 2004. Testing agency theory with entrepreneur effort and wealth. J. Finance 60, 539–576.

Blackwell, D.W., Winters, D.B., 1997. Banking relationships and the effect of monitoring on loan pricing. J. Finan. Res. 20, 275–289.

Bodenhorn, H., 2003. Short-term loans and long-term relationships: Relationship lending in early America. J. Money, Credit, Banking 35, 485–505.

Boot, A., Thakor, A., 1994. Moral hazard and secured lending in an infinitely repeated credit market game. Int. Econ. Rev. 35, 899–920.
Boot, A., Thakor, A., Udell, G., 1991. Secured lending and default risk: Equilibrium analysis and policy implications and empirical results. Econ. J. 101, 458–472.
Bowden, R., Turkington, D., 1984. Instrumental Variables. Econom. Soc. Monogr. Quant. Econom., vol. 8. Cambridge Univ. Press, Cambridge.
Brick, I., Ravid, A., 1985. The relevance of debt maturity structure. J. Finance 40, 1423–1437.
Brick, I., Ravid, A., 1991. Interest rate uncertainty and the optimal debt maturity. J. Finan. Quant. Anal. 26, 63–81.
Brueckner, J., 1994. Borrower mobility, adverse selection, and mortgage points. J. Finan. Intermediation 3, 416–441.
Canovas, G.H., Solano, P.M., 2003. Bank Relationships: Effects on the Debt Terms of the Small Spanish Firms. Polytechnic University of Cartagena, Cartagena.
Carey, M., Post, M., Sharpte, S., 1998. Does corporate lending by banks and finance companies differ? Evidence on specialization in private debt contracting. J. Finance 53, 845–878.
Chan, Y., Kanatas, G., 1985. Asymmetric valuation and the role of collateral in loan agreements. J. Money, Credit, Banking 17, 85–95.
Chakraborty, A., Hu, C., 2006. Lending relationships in line-of-credit and nonline-of-credit loans: Evidence from collateral use in small business. J. Finan. Intermediation 15, 86–107.
Chao, J., Swanson, N., 2003. Consistent estimation with a large number of weak instruments. Discussion paper No. 1417. Yale University, Cowles Foundation.
Cole, R., 1998. The importance of relationships to the availability of credit. J. Banking Finance 22, 959–977.
Cole, R.A., Goldberge, L.G., White, L.J., 2004. Cookie-cutter versus character: The micro structure of small business lending by large and small banks. J. Finan. Quant. Anal. 39, 227–251.
Conigliani, C., Ferri, G., Generale, A., 1997. The impact of bank–firm relations on the propagation of monetary policy squeezes: An empirical assessment for Italy. Banca Naz. Lavoro Quart. Rev. 202, 271–299.
Cosci, S., Meliciani, V., 2002. Multiple banking relationships: Evidence from the Italian experience. Manchester Sch. Suppl., 37–54.
D'Auria, C., Foglia, A., Reedtz, P.M., 1999. Bank interest rates and credit relationships in Italy. J. Banking Finance 23, 1067–1093.
Degryse, H., Cayseele, P., 2000. Relationship lending within a bank-based system: Evidence from European small business data. J. Finan. Intermediation 9, 90–109.
Degryse, H., Ongena, S., 2005. Distance, lending relationships, and competition. J. Finance 60, 231–266.
Dennis, S., Nandy, D., Sharpe, I., 2000. The determinants of contract terms in bank revolving credit agreements. J. Finan. Quant. Anal. 50, 87–110.
Diamond, D., 1984. Financial intermediations and delegated monitoring. Rev. Econ. Stud. 51, 393–414.
Diamond, D., 1991. Monitoring and reputation: The choice between bank loans and directly placed debt. J. Polit. Economy 99, 688–721.
Dietsch, M., 2003. Financing small businesses in France. Europ. Invest. Bank Pap. 8, 93–119.
Dunn, K., McConnell, J., 1981. A comparison of alternative models for pricing GNMA mortgage-backed securities. J. Finance 36, 471–483.
Dunn, K., Spatt, C., 1985. An analysis of mortgage contracting: implications for mortgage contract terms and pricing. J. Finance 40, 293–308.
Elsas, R., 2005. Empirical determinants of relationship lending. J. Finan. Intermediation 14, 32–57.
Elsas, R., Krahnen, J., 1998. Is relationship lending special? Evidence from credit-file data in Germany. J. Banking Finance 22, 1283–1316.
Elsas, R., Krahnen, J.P., 2002. Collateral, relationship lending, and financial distress: An empirical study on financial contracting, Unpublished manuscript. Center for Financial Studies, Frankfurt.
Ewert, R., Schenk, G., Szczesny, A., 2000. Determinants of bank lending performance in Germany. Schmalenbach Bus. Rev. 52, 344–362.
Fama, E., 1985. What's different about banks? J. Monet. Econ. 15, 29–36.
Ferri, G., Messori, M., 2000. Bank–firm relationships and allocative efficiency in Northeastern and Central Italy and in the South. J. Banking Finance 24, 1067–1095.
Gujarati, D.N., 2003. Basic Econometrics. McGraw–Hill, New York.
Guiso, L., 2003. Small business finance in Italy. Europ. Invest. Bank Pap. 8, 121–147.
Hao, L., 2003. Bank effects and the determinants of loan yield spreads. Unpublished manuscript. York University, Toronto.
Harhoff, D., Korting, T., 1998. Lending relationships in Germany: Empirical results from survey data. J. Banking Finance 22, 1317–1354.

Hester, D., 1979. Customer relationships and terms of loans: Evidence from a pilot survey: A note. J. Money, Credit, Banking 11, 349–357.

Hodgman, D., 1963. Commercial bank loan and investment policy. Bureau of Economic and Business Research, University of Illinois, Champaign, IL.

Hubbard, R.G., Kuttner, K., Palia, D., 2002. Are there bank effects in borrower cost of funds? Evidence from matched sample of borrowers and banks. J. Bus. 75, 559–581.

James, C., 1987. Some evidence on the uniqueness of bank loans. J. Finan. Econ. 19, 217–235.

John, K., Lynch, A., Puri, M., 2003. Credit ratings, collateral, and loan characteristics: Implications for yield. J. Bus. 76, 371–409.

Kane, E.J., Malkiel, B.G., 1965. Bank portfolio allocation, deposit variability, and the availability doctrine. Quart. J. Econ. 79, 113–134.

Kashyap, A., Stein, J.C., 2000. What do a million banks have to say about the transmission of monetary policy? Amer. Econ. Rev. 90, 407–428.

Kashyap, A., Stein, J.C., Wilcox, D.W., 1993. Monetary policy and credit conditions: Evidence from the composition of external finance. Amer. Econ. Rev. 83, 78–98.

Lehmann, E., Neuberger, D., 2001. Do lending relationships matter? Evidence from bank survey data in Germany. J. Econ. Behav. Organ. 45, 339–359.

Lehmann, E., Neuberger, D., Rathke, S., 2004. Lending to small and medium-sized firms: Is there an east–west gap in Germany? Small Bus. Econ. 23, 23–39.

LeRoy, S., 1996. Mortgage valuation under optimal prepayment. Rev. Finan. Stud. 9, 817–844.

Lummer, S., McConnell, J., 1989. Further evidence on the bank lending process and the capital-market response to bank loan agreements. J. Finan. Econ. 25, 99–122.

Machauer, A., Weber, M., 1998. Bank behavior based on internal credit ratings of borrowers. J. Banking Finance 22, 1355–1383.

Mallett, T., Sen, A., 2001. Does local competition impact interest rates charged on small business loans? Empirical evidence from Canada. Rev. Ind. Organ. 19, 437–452.

Menkhoff, L., Suwanaporn, C., 2003. The rationale of bank lending in pre-crisis Thailand. Unpublished manuscript. Center for Development Research ZEF, Bonn, Germany.

Miarka, T., 1999. The recent economic role of bank-firm relationships in Japan. Wissenschaftszentrum Berlin, Germany.

Orgler, Y., 1970. A credit scoring model for commercial loans. J. Money, Credit, Banking 2, 435–445.

Padilla, A., Pagano, M., 1997. Endogenous communication among lenders and entrepreneurial incentives. Rev. Finan. Stud. 10, 205–236.

Petersen, M., Rajan, R., 1994. The benefits of lending relationships: Evidence from small business data. J. Finance 49, 3–37.

Petersen, M., Rajan, R., 1995. The effect of credit market competition on lending relationships. Quart. J. Econ. 110, 407–443.

Pozzolo, A.F., 2004. The role of guarantees in bank lending. EFMA 2004 Basel Meetings. Paper available at http://ssrn.com/abstract=498982.

Rajan, R., 1992. Insiders and outsiders: the choice between informed and arm's length debt. J. Finance 47, 1367–1400.

Rajan, R., Winton, A., 1995. Covenants and collateral as incentives to monitor. J. Finance 50, 1113–1146.

Ramakrishnan, S., Thakor, A., 1984. Information reliability and a theory of financial intermediation. Rev. Econ. Stud. 51, 415–432.

Repetto, A., Rodriguez, S., Valdes, R.O., 2002. Bank lending and relationship banking: Evidence from Chilean firms. Unpublished manuscript. Universidad de Chile, Santiago de Chile.

Sapienza, P., 2002. The effect of banking mergers on loan contracts. J. Finance 57, 329–368.

Scott, J.A., Dunkelberg, W.C., 2003, A note on loan search and banking relationships. Unpublished manuscript. Temple University, Philadelphia, PA.

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

Slovin, M.B., Shushka, M.F., Polonchek, J.A., 1993. The value of bank durability: Borrowers and stakeholders. J. Finance 48, 247–266.

Streb, J.M., Bolzico, J., Druck, P., Henke, A., Rutman, J., Escudero, W.S., 2002. Bank relationships: Effect on the availability and marginal cost of credit for firms in Argentina. Unpublished manuscript. Universidad del CEMA, Buenos Aires, Argentina.

Udell, G., Berger A., 2005. A more complete conceptual framework for financing of small and medium enterprises. Unpublished manuscript. The World Bank, Policy Research Working Paper Series: 3795.

Uzzi, B., 1999. Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. Amer. Soc. Rev. 64, 481–505.
von Thadden, E., 2004. Asymmetric information, bank lending, and implicit contracts: The winner's curse. Finance Res. Letters 1, 11–23.
Weinstein, D.E., Yafeh, Y., 1998. On the costs of a bank centered financial system: Evidence from the changing main bank relations in Japan. J. Finance 53, 635–672.
Ziane, Y., 2004. Nombre de banques et relations de credit: Une approche empirique. Revue Econ. 55, 419–428.

THE JOURNAL OF FINANCE • VOL. LXI, NO. 1 • FEBRUARY 2006

Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets

NICOLA CETORELLI and PHILIP E. STRAHAN*

ABSTRACT

This paper tests how competition in local U.S. banking markets affects the market structure of nonfinancial sectors. Theory offers competing hypotheses about how competition ought to influence firm entry and access to bank credit by mature firms. The empirical evidence, however, strongly supports the idea that in markets with concentrated banking, potential entrants face greater difficulty gaining access to credit than in markets in which banking is more competitive.

ECONOMIC RESEARCH HAS FOCUSED INTENSELY in recent years on the role played by financial markets for real economic activity. Scholars have provided robust empirical evidence that broader, deeper financial markets are strongly associated, causally, with better prospects for future economic growth.[1]

Having established this basic finding, the research effort has turned to the analysis of the *mechanisms* through which finance affects real economic activity. What are the specific characteristics of financial markets that seem to affect firms and industries in nonfinancial sectors of production?[2] And, what specific characteristics of firms and industries are especially affected by finance so that it eventually translates into higher economic activity? For example, does it matter who receives credit in an industry, entrants or incumbent firms? Is it the same if an industry has a few large firms, or if instead entry always brings in younger, smaller firms?

*Cetorelli is at the Federal Reserve Bank of New York and is a fellow at the Wharton Financial Institutions Center. Strahan is at Boston College and is a fellow at both the Wharton Financial Institutions Center and the NBER. Most of the work was done while the first author was at the Federal Reserve Bank of Chicago. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Banks of Chicago and New York, or the Federal Reserve System. We thank an anonymous referee for very helpful comments. We also thank Jeremy Stein, Luigi Zingales, and seminar participants at Boston College, the Federal Reserve Banks of Chicago and New York, Ente Einaudi, Wesleyan University, and UC Davis for their inputs.

[1] This work is based on ideas tracing back at least to Schumpeter (1912), and is inspired by the early contributions of Goldsmith (1969), Gurley and Shaw (1955), and McKinnon (1973). Among the most recent contributors, we cite the work of King and Levine (1993a,b), Demirguc-Kunt and Maksimovic (1998), Levine and Zervos (1998), Rajan and Zingales (1998), and Levine, Loayza, and Beck (2000), among others.

[2] For example, Laporta, Lopez-de-Silanes, and Shleifer (2002) investigate the role of government ownership in banking; Levine (1999) focuses on the importance of the quality of legal protection of creditor rights; Jayaratne and Strahan (1996) and Cetorelli and Gambera (2001) instead explore the role of bank competition on firms' growth performance.

This paper goes straight to the heart of this line of research by investigating the impact of well-defined characteristics of banking markets on equally well-defined characteristics in product markets. More precisely, we investigate the impact of bank concentration and bank deregulation on measures of industry structure in nonfinancial sectors. We ask whether concentration of market power in banking has an effect on the number of firms in a given sector, on average firm size, and on the overall firm-size distribution.

The number of firms in a sector, average firm size, and relative proportion of small and large firms are all important factors that have a bearing on conduct and market performance. Such factors are therefore important determinants of a sector's capital accumulation and consequently of its contribution to overall economic growth. Seeking evidence of a link with specific characteristics of the banking industry thus brings the focus of analysis exactly on the question of how finance affects real economic activity.

Using data on U.S. local markets for banking and nonfinancial sectors, we find that more vigorous banking competition—that is, lower concentration and looser restrictions on geographical expansion—is associated both with more firms in operation and with a smaller average firm size. In fact, we find that the share of firms in the smallest size category (fewer than five employees) increases most dramatically with better bank competition, mostly at the expense of firms with between 100 and 1,000 employees. We find no effect of changes in banking competition, however, for the largest firms, which seems sensible given that these firms generally have access to nationwide securities markets.

Whether bank competition is "good" or "bad" for economic activity has been and continues to be a lively topic of research and policy analysis.[3] In addition to the conventional argument that concentration of market power in banking means lower equilibrium amounts of credit, it has also been claimed that banking market power is actually needed for banks to establish valuable lending relationships. Hence, whether more or less competition in banking is socially desirable is still under discussion. This paper thus contributes to expand our understanding of the economic role of bank concentration and competition.

Various related streams of literature have focused on determinants of product market competition (e.g., Brander and Lewis (1986), Chevalier (1995), Kovenock and Phillips (1995, 1997), Maksimovic (1988)), on firm size (e.g., Kumar, Rajan, and Zingales (2001), Campbell and Hopenhayn (2003)), and on firm-size distribution and more general industry dynamics (e.g., Lucas (1978), Jovanovic (1982), Evans (1987), Hopenhayn (1992)). This paper relates to these parallel lines of research and makes a contribution bridging them together.

Our evidence is consistent with the findings documented in several recent papers that focus on banking concentration and competition policies across

[3] A conference titled "Bank Competition: Good or Bad?" was organized in 2000 by The Wharton School and the Center for Financial Studies at Frankfurt University. More recently, two conferences on the role of bank concentration and competition have been organized by the World Bank and by the Cleveland Federal Reserve and the Journal of Money, Credit, and Banking.

countries. Cetorelli (2001) provides evidence of larger average firm size in countries with more concentrated banking. Along similar lines, Cetorelli (2004) finds that enhanced bank competition following passage of the Second European Banking Directive brought a reduction in average firm size. Matching data on job creation and destruction in U.S. manufacturing sectors with banking data across U.S. markets, Cetorelli (2003) shows that more bank concentration implies less entry and growth of younger firms and also delayed exit of older firms. Again based on cross-country data, Beck, Demirguc-Kunt, and Maksimovic (2004) find that higher bank concentration is associated with more financing obstacles, especially for smaller firms. Bertrand, Schoar, and Thesmar (2004) find evidence that bank deregulation in France led to more entry in bank-dependent sectors of production. Exploiting historical census data from 19$^{\text{th}}$ century Mexico, Maurer and Haber (2003) find evidence that bank concentration favored lending to "connected" borrowers in the domestic textile industry over other competitors, even though these competitors exhibited higher efficiency measures. In contrast, Bonaccorsi and Dell'Ariccia (2004) find that concentration in banking reduces entry rates for Italian firms in industries with relatively opaque assets (i.e., few intangible assets), relative to entry in industries with less asset opacity.[4]

Ours is the first study to explore not only how average firm size responds to banking competition, but how the whole size distribution responds. This is an important improvement, in that we are able to test more directly whether more or less bank competition is beneficial for all firms in a sector or whether instead, the effect may be different for firms in distinct size classes. The analysis on the size distribution will also allow us to characterize more fully the results obtained in the literature on average size. If, indeed, changes in bank competition have an impact on the average firm, we will be able to tell whether the effect is mainly on the smaller firms or on the larger ones or both. Moreover, our data allow us to measure banking structure at a more local level. This is a distinct advantage because much of the research on bank market power suggests that the relevant geographical market for banking services, especially for small firms or potential entrepreneurs, is local (see, for example, Berger, Demsetz, and Strahan (1999)).

In the remainder of the paper, we first flesh out the theoretical links between banking concentration and industrial structure in order to motivate our empirical tests. In Section II, we discuss our identification strategy, while in

[4] More broadly, financial development is positively related to firm entry. Guiso, Sapienza, and Zingales (2004a) show that regions in Italy with higher levels of social capital (measured by voter turnout or blood donations) have more developed financial markets. For example, social capital is positively related to the use of checks, is negatively associated with cash holdings by households, and is positively related to household holdings of stock. They then show in a subsequent study (Guiso, Sapienza, and Zingales (2004b)) that a greater degree of financial development in turn increases the rate of self-employment, lowers the average age of the self-employed, and increases the rate of creation of new firms. These changes in outcomes seem to be associated with faster firm-level and overall economic growth. Klapper, Laeven, and Rajan (2004) also find that financial development increases firm entry using cross-country and cross-industry data.

Section III, we present the data set and the main variables used in the analysis. Section IV documents the empirical results, and Section V concludes.

I. Theoretical Arguments

How does bank competition affect the market structure of nonfinancial industries? As pointed out in Cetorelli (2001), several countervailing forces are potentially in play. One force emphasizes that lending to opaque firms requires that the bank and the borrower forge a long-term relationship. Information gained over the course of time by the bank can be used to make value-enhancing credit decisions (i.e., expand credit to "winners" and restrict credit to "losers"). However, banks can sustain the cost of starting a relationship with unknown, risky entrepreneurs only if market power allows them to recoup the cost at later stages if such entrants turn out to be successful. This idea was first presented by Mayer (1988) and then formalized and tested by Petersen and Rajan (1995).[5] To the extent that it forecloses the opportunity to extract profits from successful relationships, vigorous competition may mitigate willingness of banks to invest in relationships at all. This force, applied to our case, suggests that banks with market power should guarantee more industry entry than competitive banks. At the same time, because banks with market power charge higher rates to the older firms than would competitive banks, firms in noncompetitive banking markets may not grow as large as they could otherwise. All together, and ceteris paribus, one should expect to find more firms in an industry, a lower average firm size, and a larger prevalence of small rather than large firms where banks have more market power.

Two countervailing forces suggest that market power may both dampen entry and reduce the relative importance of smaller firms. First, bank market power may reduce credit availability generally. This standard channel, whereby increased concentration in banking leads to less credit supply and higher loan prices, justifies antitrust enforcement. While less credit hurts all firms, smaller firms and potential entrants are likely to be more reliant on bank credit than larger and better established firms. Thus, these smaller firms may be harmed more by reduced credit supply than larger firms.

In addition to this standard channel, banks with market power may tend to favor their established borrowers over new borrowers. The value of a bank's current lending relationships will depend on the future profitability of its borrowers, which in turn depends on prospective entry and growth of new competitors. A bank's incentive to support the profitability of its older clients could thus restrain its willingness to extend credit to potential industry entrants (or emerging small firms). In recent papers, Spagnolo (2000) and Cestone and White (2003) present theoretical frameworks in which existing lending relationships do indeed affect the behavior of lenders vis-à-vis potential new

[5] Another solution is for the lender to hold an equity or equity-like claim against the establishment, as is commonly observed in the venture capital industry.

borrowers.[6] The less competitive the conditions in the credit market, the lower the incentive for lenders to finance newcomers. Hence, banking concentration (as well as regulatory impediments to competition) can represent a form of financial barrier to entry in product markets. Banking market power may lead to fewer firms, a larger average firm size, and a higher proportion of large firms in markets in which banks have more market power.[7]

II. Identification Strategy

These ideas suggest that banking competition potentially plays a key role in determining industry structure in product markets. Either prospective entrants benefit at the expense of older incumbents with more bank competition or vice versa.

The difficulty in empirical implementation is that there may be common factors that drive the structure of both banking and industrial sectors that are difficult to measure and thus control in a regression. For example, small markets may have fewer banks and also fewer nonfinancial firms. Finding a positive association between bank concentration and product market concentration could simply pick up this market characteristic. By the same token, banks have increasingly turned to analytical credit scoring models based on systematic statistical processing of information about borrower credit history, financial strength, and asset values. This growing importance of "hard information" may have led to a corresponding decline in the value of lending relationships over time. If technology adoption (which is difficult to measure empirically) exerts an independent effect on the structure of both banking and nonfinancial industry, then our ability to sort out how bank concentration affects industry structure may be confounded.[8]

Given these concerns, identification based on direct effects of changes in bank competition across industries in a market will not take us far. We proceed instead with the idea that banks play a more important role for firms in some industries than in others. Rajan and Zingales (1998) point out that industries differ in their dependence on external sources of finance. These differences stem from industry-specific technological factors affecting initial project scale, gestation period, cash-harvest period, and needs for further investments (Rajan and Zingales (1998), p. 563). So, for example, a firm in a sector such as Tobacco

[6] Also related are the contributions of Bhattacharya and Chiesa (1995) and Hellmann and Da Rin (2002).

[7] This effect should be especially strong where the boundaries of product markets and lending markets coincide. Some product markets in our data set may span multiple banking markets. In such cases, a bank could still have an impact on entry within its area of influence. Because we are unable to define product market boundaries, however, our results are, if anything, biased toward finding no result.

[8] Information technology may also have played a role in explaining deregulation of banking, although Kroszner and Strahan (1999) find that the strength of interest groups can explain the precise timing of deregulation across states.

or Leather is less likely to depend on external sources of finance (such as bank lending) than firms producing Electronics or Chemical Equipment.

Given the richness of our data set (to be illustrated in the next section), we pursue identification by exploiting cross-industry differences in external financial dependence. Firms in sectors more dependent on external finance ought to be affected more by variation (both across time and across markets) in bank competition. More precisely, we use nondependent sectors as a control group, which in an "experiment" designed to gauge the effect of a change in bank competition is not expected to exhibit a response. We can instead use dependent sectors as a treatment group that in the same experiment is expected to exhibit a response. We thus achieve identification by measuring the differential effect of a change in bank competition between the two groups. Going back to the first of our previous examples, if none of the theoretical arguments presented earlier bear any justification in the data, and if market size is all there is behind a positive correlation between bank concentration and product market structure, then a change in bank competition should not have any differential effect on treatment and control groups. On the other hand, if bank competition actually implies greater opportunities for entry, then such effects should be larger on the treatment sectors.

We implement this strategy by estimating regression models that fit the following general structure:

$$\begin{aligned} Y_{jst} &= \alpha \cdot \mathit{Employment\ share}_{jst} + \delta \cdot \mathit{Market\ trends}_{st} + \gamma \cdot \mathit{Industry\ trends}_{jt} \\ &\quad + \beta \cdot \mathit{Bank\ dependence}_{j} \cdot \mathit{Bank\ competition}_{st} + \varepsilon_{jst}, \end{aligned} \tag{1}$$

where in separate regressions Y equals the number of firms, average firm size, or a measure of the entire size distribution of industry sector j in market s over time t. Industry j's share of total manufacturing employment in market s over time t ($\mathit{Employment\ share}_{jst}$) controls for the relative importance of a given sector in a market.[9] The $\mathit{Market\ trends}_{st}$ are vectors of market/time indicator variables controlling for any local market, time-varying effect on industry structure. Similarly, $\mathit{Industry\ trends}_{jt}$ are industry/time indicator variables controlling for industry-specific, time-varying factors. By including the two sets of fixed effects, we fully absorb any common confounding factor that could trouble identification, thereby minimizing the risk that our results are driven by either reverse causality (changes in industry structure driving changes in banking structure), or an omitted factor (e.g., technological changes that drive both banking structure or regulations and the structure of nonfinancial businesses). Consistent with the identification strategy, note that the direct effect of

[9] In studies of cross-sector industrial growth, such a measure consistently predicts that sectors that had grown substantially in the past, and therefore are already relatively large, grow less in the future (see Rajan and Zingales (1998) and Cetorelli and Gambera (2001)). Moreover, theories of an industry's lifecycle predict that a sector that has already grown substantially should experience less intensive firm entry (see Klepper (1996)). Hence, the share variable controls for the stage in which a sector is within its industry, and specifically it should capture the different intensity in entry due to lifecycle-specific reasons.

Bank dependence$_j$ is not identified because its effects are fully captured by the set of industry-level fixed effects. Similarly, the direct effect of *Bank competition$_{st}$* is also not identified; its effects are fully captured by the market-level fixed effects.

The coefficients of interest, which measure the effect of bank competition, are identified by estimating a vector of three separate interaction terms between industry-specific external bank dependence measures and three alternative measures of bank competition (described below). For instance, consider a regression in which the dependent variable is a measure of the number of firms in operation in each industry. If an increase in bank competition facilitates entry, then this increase in the number of firms should be especially pronounced for sectors with a high degree of bank dependence. Thus, the interaction of bank dependence with bank competition will be positive and significant. The opposite would be true if entry were more likely where banks have market power.

III. Variable Definitions and Data Sources

We construct a panel data set of manufacturing establishments in operation across U.S. states between 1977 and 1994.[10] Data on establishments are available at a disaggregated level on an annual basis from the County Business Patterns, which is an annual survey by the Census Bureau. These data provide the best way to consider industry structure over a long span of time at a disaggregated level. Moving to a more finely disaggregated level, either by industry SIC code or by locality, creates substantial difficulties with missing values, so we focus on the two-digit level of aggregation by industry and the state level for geography. We focus only on industries within the manufacturing sector. From this data set, we compute the total number of establishments in an industry/state/year and the average establishment size (workers per establishment). As shown in Table I, there are, on average, 0.07 establishments per capita, and the average establishment has 69 workers.

To characterize the whole distribution of establishment sizes, we construct the share of establishments in an industry, state, and year in each of the five categories: establishments with fewer than 5 employees, establishments with 5–19 employees, establishments with 20–99 employees, establishments with 100–999 employees, and establishments with 1,000 or more employees. Unconditionally, 31% of establishments are in the smallest size category, while 86% of establishments have fewer than 100 employees (Table I).

Before moving forward, it is worth noting that our data are based on employment at establishments rather than firms. An establishment is an economic unit in which production occurs, such as a plant, a factory, or a restaurant that employs people. Thus, there is some measurement error in our dependent variable induced by the fact that large firms often own many establishments. Data from *The Economic Census*, however, suggest that most firms are composed of

[10] We justify the choice for the sample period when we describe the history of deregulation of the U.S. banking industry.

Table I
Summary Statistics

This table reports the summary statistics for data used in our panel regressions. The unit of observation varies at the state-industry-year level. Data on establishments are from the County Business Patterns, 1977–1994. The Herfindahl–Hirschman Index (HHI), our measure of concentration, is built from HHIs at the city level (Metropolitan Statistical Area (MSA)). By city, we construct the sum of squared share of deposits for each banking company. We then average this concentration measure across all MSAs within the state, weighted by total deposits in each MSA (source: FDIC Summary of Deposits). The post-branching deregulation indicator equals one during all years in which a state permits in-state branching; the post-interstate banking deregulation indicator equals one during all years in which a state permits out-of-state banking companies to buy banks headquartered in the state. For a detailed discussion of the banking deregulation, see Kroszner and Strahan (1999).

	Mean	Standard Deviation
Establishments per capita	0.07	0.09
Share of establishments with < 5 employees	0.31	0.15
Share of establishments with 5–19 employees	0.31	0.12
Share of establishments with 20–99 employees	0.24	0.11
Share of establishments with 100–999 employees	0.13	0.12
Share of establishments with more than 1,000 employees	0.01	0.02
Average establishment size (employees per establishment)	69	83
HHI (sum of squared local market deposit share)	0.19	0.07
Post-branching deregulation indicator	0.60	–
Post-interstate banking deregulation indicator	0.45	–

just one establishment. For example, in 1997 there were a total of 316,952 manufacturing companies; of these, 295,876 were single-establishment companies, about 94% of the total. Also, at the same time there were a total of 381,696 establishments in operation. Therefore, in 1997, single-establishment companies represented 78% of total establishments.[11]

A. Constructing an Instrument for Bank Dependence

As we discuss in the previous section, the effects of banking competition on industry structure should depend on the relative bank dependence of firms in an industry. The trick empirically is to construct a measure of bank dependence that reflects demand for bank finance, rather than one that confounds demand-side effects with variation in the availability of credit supply from banks.[12] As indicated by the summary statistics for establishments, most firms are small.

[11] These proportions are similar looking back at previous census data. Early research shows, for example, that the rate of creation of new businesses is correlated with the share of new establishments in a local economy (Black and Strahan (2002)). The existence of a close correlation between the number of establishments and the number of firms is also documented in Cetorelli (2001) for a cross-section of countries.

[12] Variation in bank credit supply introduces noise (measurement error) into the actual use by small firms of bank finance. Moreover, the extent of that noise will be greater for firms in industries that are more bank dependent.

Moreover, the theoretical arguments illustrated earlier all emphasize the role of banks for small firms and potential entrants. We thus begin by constructing a measure of bank finance using data from the 1998 *Survey of Small Business Finance* (*SSBF*). This survey was conducted by the Federal Reserve and covers a sample of 3,561 small firms with fewer than 500 employees. The sample was designed to be nationally representative but was structured to ensure representation across firm-size categories, location, and race of the owner. For each two-digit SIC manufacturing sector, we construct firms' share of assets financed with debt (loans, capital leases, and lines of credit) from financial institutions for the firm at the median of the distribution. These "loans" are supplied mainly by commercial banks (70% of the surveyed firms use banks for credit), but they also include some funds from other depository institutions (thrifts, credit unions) as well as unregulated finance companies (Bitler, Robb, and Wolken (2001)). This variable represents the actual use of bank finance by small firms, as a share of their balance sheet.[13]

Despite the appeal of such a measure to capture a firm's need for bank finance, the very fact that the data is based on small firms, with relatively limited access to external suppliers of finance, raises the possibility that such accounting measures may not be clean measures of demand. We therefore proceed by constructing an instrument for bank dependence using information from Compustat firms and justify that choice by documenting its high correlation with small firms' actual use of bank (and other intermediary) funds. Our procedure follows closely the one described in Rajan and Zingales (1998). The key identifying assumption, as in Rajan and Zingales, is that the use of finance by Compustat firms allows us to observe their demand for external funds. These firms are large and well established, with far better access to well-developed U.S. securities markets than small firms. Hence, there is a much higher likelihood that observed financial policy is not skewed by constraints on the supply side.

We begin by taking all of the Compustat firms between 1980 and 1997 and carefully separate firms based on the number of years they have been on Compustat. More precisely, we exclude "young" Compustat firms, that is, firms that had gone public only recently. A measure of external financial dependence for such firms is likely to be nonrepresentative of the typical needs of a firm in a given sector due to selection bias: firms that have just gone public have a voracious appetite for external funds—satisfying this high demand is a large reason that they go public in the first place. Following Rajan and Zingales (1998), we define as young those firms in Compustat for 10 years or less, and we define as "mature" those firms appearing in Compustat for more than 10 years.[14]

[13] The *SSBF* only reports data for a single year, hence balance sheet measures are more representative of financial policy than flow-based measures using capital expenditures and gross cash flow.

[14] We also change the cut off point to 5 years instead of 10. The resulting new measure of external financial dependence for "mature" firms has a correlation of 0.98 with that used following Rajan and Zingales (1998). The regression results that we present later are not qualitatively affected by the choice of this alternative definition of mature firms.

To construct external financial dependence, we sum across all years each firm's total capital expenditures (Compustat item #128) minus cash flow from operations. Cash flow from operations equals revenues minus nondepreciation costs (Compustat item #110) plus decreases in inventories and accounts receivable plus increases in accounts payable.[15] This sum equals the total external funds needed to finance investments. If the total is negative, it means that the firm had free cash flow available for disbursement to shareholders or to pay down debt; otherwise, the firm needed to raise additional capital to finance its investment. We then divide this free cash flow figure by total capital expenditure. After constructing this ratio for each firm, we use the median value for all firms in each two-digit SIC category.

Table II reports our measures of external financial dependence for mature Compustat firms, and that calculated from the 1998 SSBF. Looking across sectors, we find that leather and leather products, tobacco manufactures, and apparel have the lowest need for external funds as mature firms, whereas electronic equipment and chemicals and related products exhibit the highest need for external finance. Moreover, the correlation between the two measures of external financial dependence is high ($\rho = 0.51$), thus suggesting that external financial dependence for mature Compustat firms makes a powerful instrument for small firms' demand for bank credit.

As in the previous section, we want to emphasize the distinction between low-dependent sectors (the control group) and high-dependent ones (the treatment group). Hence, we use an indicator variable for a firm's external financial dependence in our base specification that is equal to one if the use of external funds for mature Compustat firms is positive, and zero otherwise. Beyond permitting a clean interpretation of the results, the indicator for external financial dependence is not skewed by the very high level of cash flow relative to investment for three sectors—leather, tobacco, and apparel—documented in Table II. For robustness, however, we also report results using the continuous measure of external financial dependence for mature Compustat firms (rather than a zero-one indicator variable), as well as the loans-to-assets ratio for small firms.[16]

[15] In other words, we subtract investments in net working capital from cash flow. The numerator of external financial dependence equals the negative of "free cash flow" as traditionally defined in capital budgeting problems. We also construct an external financial dependence measure that normalizes this figure by net long-term investment (capital expenditures minus depreciation) and also by net total investment (i.e., capital expenditure minus depreciation plus the change in net working capital). These measures are highly correlated with the one used here ($\rho = 0.89$ and $\rho = 0.70$, respectively). Note also that the Compustat items mentioned in the text are only defined for cash flow statements with codes 1, 2, or 3. For format code 7, we use the sum of items #123, 125, 126, 106, 213, and 217.

[16] While mature firms in Compustat clearly have access to alternative sources of finance beyond banks (e.g., bonds or public equity), our identifying assumption is that demand for external credit for small sized and medium-sized firms will look similar to demand for external funds by mature Compustat firms. Since smaller firms do not have access to public securities markets, we argue that their ability to raise external funds depends critically on the competitiveness of the local banking market.

Table II

External Financial Dependence for Manufacturing Sectors

External financial dependence equals the proportion of capital expenditures financed with external funds. A negative value indicates that firms have free cash flow, whereas a positive value indicates that firms must issue debt or equity to finance their investment. The figures represent the median value for Compustat firms in each industry sector over the 1980 to 1997 period. Mature firms are those that have been on Compustat for 10 years or more. The rows are sorted by the external finance measure for mature firms, which is our baseline measure of an industry's long-term financing needs. The loans/assets ratio is the median ratio of loans to assets for small firms from the Federal Reserve 1998 Survey of Small Business Finance.

Industry Sector (Two-Digit SIC)	Median Loans/Assets for 1998 SSBF Firms	External Financial Dependence for Mature Compustat Firms
Leather and leather products (31)	0.04	−0.96
Tobacco manufactures (21)	N/A	−0.92
Apparel and other textiles (23)	0.13	−0.61
Food and kindred products (20)	0.12	−0.24
Fabricated metal products (34)	0.27	−0.24
Furniture and fixtures (25)	0.36	−0.23
Stone, clay, glass, and concrete products (32)	0.31	−0.20
Miscellaneous manufacturing (39)	0.28	−0.20
Printing and publishing (27)	0.33	−0.07
Instruments and related products (38)	0.29	−0.04
Transportation equipment (37)	0.06	0.01
Industrial machinery and equipment (35)	0.21	0.01
Primary metal industries (33)	0.31	0.03
Lumber and wood products (24)	0.49	0.04
Rubber and plastic products (30)	0.30	0.04
Paper and allied products (26)	0.37	0.06
Petroleum and coal products (29)	0.60	0.09
Textile mill products (22)	0.47	0.10
Electrical and electronic equipment (36)	0.14	0.22
Chemicals and allied products (28)	0.33	0.28
Median	0.30	0
Correlation between the two	0.51	

Table III compares investment and financing behavior for Compustat firms by our measure of external financial dependence. These simple comparisons show that differences in external financial dependence are not correlated with differences in investment rates, market valuation, or size. For both groups, for example, the investment rate (capital expenditures divided by property, plant, and equipment), the book-to-market equity ratio, and real sales are approximately equal. Also, external funds—for those firms that continue to need them—are raised mostly by borrowing; the median mature firm uses no equity to finance new projects. Thus, the greater demand for external finance at mature firms does not seem to reflect greater growth or investment opportunities; instead, as suggested by Rajan and Zingales (1998), external finance reflects differences in the incubation period before assets begin to generate

Table III

Median Industry-Level Investment and Financing Characteristics for Mature Compustat Firms, by External Financial Dependence

This table reports the average of the industry-level median characteristics for mature Compustat firms (i.e., firms in Compustat for at least 10 years). External financial dependence equals the share of capital expenditures financed with external funds; equity finance equals the share of capital expenditures financed by issuance of new equity (negative values indicate that cash is paid out to equity holders), and debt finance equals the share of capital expenditures financed by new borrowing (negative values indicate reductions in outstanding debt). Real sales are expressed in 1997 dollars, based on the Consumer Price Index.

	High External Financial Dependence for Mature Firms	Low External Financial Dependence for Mature Firms
External financial dependence	0.05	−0.23
Equity finance	0.00	0.00
Debt finance	0.05	−0.23
Capital expenditures/property, plant and equipment	0.19	0.19
Ratio of book value of equity to market value of equity	0.66	0.68
Median real sales (1997 millions of dollars)	\$273	\$216

sufficient cash to finance the firm. Thus, we can be confident that differences in response to bank competition across these two groups of industries reflect differences in their financing needs, rather than differences in their real investment behavior.[17]

B. Competition in the Local Banking Market

We focus on several measures of competition in the local banking industry. Our first two measures exploit policy innovations. Restrictions on bank expansion across geographical boundaries in the United States date back to the 19th century. Although there was some deregulation of branching restrictions in the 1930s, most states either prohibited branching altogether (the "unit banking" states) or limited branching until the 1970s, when only 12 states allowed unrestricted statewide branching. Between 1970 and 1994, however, 38 states deregulated their restrictions on branching (see Jayaratne and Strahan (1996), Kroszner and Strahan (1999), and Stiroh and Strahan (2003)).

In addition to facing restrictions on in-state branching, the Douglas Amendment to the 1956 Bank Holding Company Act prohibited a bank holding company from acquiring banks outside the state in which it was headquartered unless the target bank's state permitted such acquisitions. Since no state

[17] Implicit in our identification strategy is the assumption that external financial dependence is constant over time, or at least that the industry ordering is not altered substantially. Since we compute external financial dependence using mature Compustat firms, presumably closer to industry steady state conditions, this assumption seems reasonable.

allowed such transactions in 1956, the amendment effectively barred interstate banking organizations. Starting in the early 1980s, many states began to enter regional or national reciprocal arrangements whereby their banks could be bought by any other state in the arrangement. This history presents us with a convenient way to test how industry structure in nonfinancial sectors has been affected by the increased competition (real and potential) in banking that followed state-level deregulation.[18]

We capture the effects of each type of deregulation by including an indicator variable equal to one after a state permits branching by means of merger and acquisition within its borders, and another indicator variable equal to one after a state permits interstate banking (that is, after a state allows bank holding companies from other states to buy its banks).[19] The two types of deregulation are somewhat distinct in their effects. Deregulation of restrictions on branching reduces entry barriers into new markets and also enhances the corporate takeover market by making it easier for banks to gain control over other banks' assets. With full branching deregulation, a bank may enter a new market, either by buying existing branches or by opening new branches. Also, the cost of acquiring another bank is reduced because an acquiring bank may merge the target bank's operation into its existing franchise. By reducing entry barriers, branching deregulation constrains banks from exploiting market power. Interstate banking deregulation, however, only affects who can own bank assets: Prior to deregulation, only bank holding companies located within a state could buy banks chartered in that state, while after deregulation, bank holding companies operating in other states could also do so.

In addition to looking at changes in competition induced by deregulation of the industry, we also include a direct measure of local market concentration, equal to the deposit Herfindahl–Hirschmann Index (HHI). The HHI is calculated as the deposit-weighted average of the HHI indexes of the Metropolitan Statistical Areas (MSAs) in a state/year. The HHI for each local market is defined as the sum of squared market shares, where market shares are based on branch-level deposit data from the Federal Deposit Insurance Corporation's (FDIC) Summary of Deposits data set.[20] So, for example, if a bank owned

[18] Early research by Peltzman (1969) suggests that the balkanized U.S. banking system affected the transmission of monetary policy, whereby deposits at banks that were permitted to branch responded more quickly to changes in bank reserves (controlled by the Federal Reserve) than deposits at unit banks. Deregulation of restrictions on bank expansion, both within and across states, has been shown to improve bank efficiency, enhance corporate control, and limit market power. See Jayaratne and Strahan (1998).

[19] Most states first permit banks to branch by buying existing branches in new markets or by purchasing whole banks and then creating branches out of the purchased bank's offices. Then, states typically open up their markets to unrestricted branching such that banks may open new branches anywhere in the state.

[20] The deposit HHI is the standard tool used in antitrust oversight of bank mergers. Local markets (usually MSAs or non-MSA counties) with HHI below 0.18 are deemed to be served by enough banks to assume that conditions are competitive. For localities with HHI above 0.18, antitrust concerns by the Federal Reserve and the Department of Justice are sometimes raised. See Berger, Demsetz, and Strahan (1999) for an overview of bank mergers and antitrust policy.

10 branches within an MSA, this bank's market share would equal the sum of all of its deposits in those 10 branches, divided by the total deposits held by all bank branches within that market. For a market with a single bank owning all of the branches, the HHI would equal one, whereas in a perfectly atomistic market the HHI would approach zero.

The history of U.S. banking deregulation defines the lower and upper bounds of our sample period. Our panel data starts in 1977, the beginning of the period of dramatic state-level deregulation. We end the sample in 1994, when deregulation of restrictions on the ability of banks to expand across local markets was completed with the passage of the Riegle–Neal Interstate Banking and Branching Efficiency Act. After 1994, it becomes increasingly less plausible to view markets in banking as local, both because of the completion of deregulation and because the advent of new technologies in bank lending began to allow banks to lend to borrowers not physically close to their bank. For example, Petersen and Rajan (2002) show that banks during the 1990s are much more likely to lend over long geographic distances than they were in the 1970s. Also, banks began to operate branches across state lines for the first time in 1995.[21]

IV. Results

A. *Difference-in-Differences Estimates*

As noted above, our identification strategy is to emphasize the differential effects of bank competition across bank-dependent and bank-nondependent sectors. Table IV illustrates this strategy based on a simple "difference-in-differences" estimate of the effect of changes in concentration and regulation on establishment size. We simply average the data across four groups: sectors with high and low external financial dependence and markets/years, with high and low bank concentration (Panel A) or with tight and relaxed bank regulation (Panel B). Establishment size is consistently higher among sectors with high external financial dependence (the treatment group) than sectors with low external financial dependence (the control group). However, our estimate depends on the relative response of the two groups to changes in competition. For example, in the treatment group establishment size is 11 employees higher in banking markets with low concentration (HHI below 0.14) compared to highly concentrated markets (HHI above 0.24). On the other hand, establishment size in the control group is 20 employees higher in the less concentrated banking markets. Thus, the estimated effect of decreased banking concentration (increased competition) is a decline of nine employees on average: $11 - 20 = -9$.

Panel A thus illustrates the importance of our identification strategy; by focusing on the direct effect of concentration, one might conclude that raising banking competitiveness leads to larger firms. Establishment size, however, is negatively correlated with banking market concentration for reasons having

[21] For example, NationsBank consolidated banks from several other states into its primary North Carolina bank (NationBank NC N.A.), leading to an increase of this bank's assets from \$31 billion in 1994 to \$79 billion in 1995.

Table IV

Comparison of Mean Establishment Size for Competitive versus Noncompetitive Banking Markets

This table reports a simple difference-in-differences estimate of the effects on average establishment size of moving from a regulated to an unregulated market or from a concentrated to an unconcentrated market. Concentrated markets have a Herfindahl–Hirschman Index (HHI) in the first quartile of the distribution (HHI > 0.24); unconcentrated markets have an HHI in the top quartile (HHI < 0.14). Regulated markets allow neither branching nor interstate banking; deregulated markets permit both. The treatment group is represented by industries with above-median needs for external finance. The control group is represented by industries with below-median needs for external finance. * denotes statistical significance at the 5% level; and ** denotes significance at the 1% level, based on a simple *t*-test.

Panel A

	Banking Markets with Low Concentration	Banking Markets with High Concentration	Difference
(1) Treatment group	82	71	11
(2) Control group	61	41	20
Difference-in-Differences =	21	30	−9**

Panel B

	Banking Markets with Deregulated Banking Markets	Banking Markets with Regulated Banking Markets	Difference
(3) Treatment group	74	90	−16
(4) Control group	52	56	−4
Difference-in-differences =	22	34	−12**

nothing to do with banking competition. Rural and relatively sparsely populated states tend to have fewer banks and hence higher measures of banking concentration, and these states also tend to have smaller establishments on average. Comparing the response of the two industry sectors removes these state effects from the difference-in-differences estimate. In the regressions below, state differences are removed by including the full set of state/year fixed effects.

In Panel B, we consider varying the regulatory status of banking markets. Moving from fully regulated (no branching or interstate banking) to fully deregulated (branching and interstate banking permitted), average establishment size declines by four employees in the control group. In contrast, size falls by 16 employees in the treatment group. In this example, then our difference-in-differences estimate of the effect of deregulation would equal a decline of 12 employees: $-16 - (-4) = -12$. These simple mean comparisons illustrate the identification strategy embedded in our regressions (see equation (1)), which we turn to next.

Table V

Regression of Log of Establishment Size and Log of Number of Establishments on Bank Competition Variables

Each column in this table reports statistics from a fixed effects regression, where the dependent variable is the log of the total number of establishments per capita (columns labeled "Number") or the log of average number of workers per establishment (columns labeled "Size"). Data on establishments are from the County Business Patterns, 1977–1994. The terms of interaction are based on external financial dependence for mature Compustat firms and the ratio of total loans to assets for small firms based on the Federal Reserve 1998 Survey of Small Business Finance (see Table II). The Herfindahl–Hirschmann Index (HHI), our measure of concentration, is built from HHIs at the city level (Metropolitan Statistical Area (MSA)). By city, we construct the sum of squared share of deposits for each banking company. We then average this concentration measure across all MSAs within the state, weighted by total deposits in each MSA (source: FDIC Summary of Deposits). The post-branching deregulation indicator equals one during all years in which a state permits in-state branching; the post-interstate banking deregulation indicator equals one during all years in which a state permits out-of-state banking companies to buy banks headquartered in the state. For a detailed discussion of the banking deregulation, see Kroszner and Strahan (1999). Industry share of employment equals the total employment in a given industry-state-year divided by the total employment in the corresponding state-year. Absolute values of t-statistics are in parentheses, where * indicates significance at the 5% level and ** indicates significance at the 1% level.

	External Financial Dependence Measure					
	Indicator = 1 If Industry Has Positive External Dependence for Mature Compustat Firms		Level of External Dependence for Mature Compustat Firms		Industry Median Loans/Assets for Firms in the 1998 SSBF	
	Number	Size	Number	Size	Number	Size
Post-branching × external financial dependence	−0.006 (0.30)	−0.018 (0.77)	−0.025 (0.81)	0.120 (3.22)**	−0.095 (1.40)	−0.145 (1.96)*
Post-interstate × external financial dependence	0.122 (4.14)**	−0.105 (2.98)**	0.039 (0.77)	−0.263 (4.41)**	0.152 (1.58)	−0.904 (8.04)**
Local market HHI × external financial dependence	−0.457 (4.03)**	0.934 (6.88)**	−1.245 (6.56)**	2.377 (10.49)**	−1.104 (2.96)**	3.490 (8.02)**
Industry share of employment	7.647 (94.44)**	6.672 (68.83)**	7.652 (94.57)**	6.675 (69.11)**	7.581 (101.56)**	6.497 (74.54)**
Observations	15,127	15,127	15,127	15,127	14,717	14,717
Fixed effects	State × Year, Industry × Year					
R^2	0.89	0.56	0.89	0.56	0.86	0.62

B. Regression Results for Size and Number of Establishments

Table V reports our regression results for the log of the number of establishments and the log of establishment size (average employees per establishment). The first two columns report the results using an indicator variable to separate firms into two groups: Those with positive external financial dependence and those with negative dependence. These results are most closely related to the difference-in-differences calculations in Table IV because we can think about a treatment group response versus a control group response. Columns 3 and 4 report the interaction effects of the continuous measure of external financial dependence for the mature Compustat firms with our three banking competition measures. The last two columns are based on interacting the loans/assets ratio for small firms with bank competition. As noted above, we include a full set of industry/year and state/year fixed effects; these capture all of the direct effects of external financial dependence (an industry characteristic) and bank competition (a market characteristic) on establishment size. Thus, only the coefficients on the interaction terms are identified in these models.

All six specifications in Table V support the idea that increased bank competition through lower banking concentration is associated with greater entry (i.e., more total establishments) and a smaller average establishment size. While somewhat less robust across our measures of external financial dependence, the results also suggest that deregulation, particularly relaxation of restrictions on interstate banking, comes with an increase in the number of firms and a reduction in firm size. This finding mirrors Black and Strahan (2002), who find that interstate banking reform was followed by an increase in the formation of new incorporations.[22]

In Table VI, we use the coefficients from above (Table V) to assess the economic importance of increased banking competition on the number and size of establishments. We report effects similar to the difference-in-differences of Table IV, but now measure percentage changes rather than absolute changes since both dependent variables are measured in logs. For example, we report the percentage change in the number and size of establishments stemming from a move from full regulation (no branching or interstate banking) to full deregulation (branching and interstate banking permitted) for financially dependent industries relative to less dependent industries. Moreover, we measure the effect of moving banking concentration from the 75th to the 25th percentile of its distribution (HHI changes from 0.24 to 0.14). These percentage changes allow a simple way to assess the economic significance of the results.

Columns 1 and 2 report economic effects for our preferred measure of external financial dependence, which simply uses an indicator variable to separate sectors into low and high external financial dependence. These results suggest that deregulation, and the associated increase in banking competition,

[22] We also test for a structural break in the link from our three bank competition measures to industry structure by interacting each measure with a post-1985 indicator variable (1985 is the midpoint of our sample). This test is motivated by the increasing use of information technology in the latter half of the sample, which may have reduced the importance of relationships in bank lending and thus altered the role of market power in shaping industry structure. We find, however, no such break; these coefficients are stable over time. The F-test for the structural break equals 1.17 (p-value = 0.31) for the model based on the number of establishments, and 1.32 (p-value = 0.26) for the model based on average establishment size. A plausible explanation for this is that the widespread adoption of the new lending technologies really only occurred during the last 10 years or so, hence outside of our sample period.

Table VI

Economic Significance of a Change in Bank Competition on Establishment Size and Number of Establishments

This table uses the coefficients from Table V to compute the effect of an increase in bank competition on the log number of establishments (log of employees per establishment) on industries that are financially dependent relative to industries that are not financially dependent. In columns 1 and 2, we report the effect of bank competition on industries with above-median external financial dependence for mature Compustat firms, relative to industries with below-median external financial dependence. The coefficients come from columns 1 and 2 of Table V. In columns 3 and 4, we report the effect of bank competition on industries with external financial dependence for mature Compustat firms at the 75th percentile of the distribution, relative to industries with external financial dependence at the 25th percentile. The coefficients come from columns 3 and 4 of Table V. In columns 5 and 6, we report the effect of bank competition on industries with loans/assets for firms in the 1998 SSBF at the 75th percentile, relative to industries at the 25th percentile. The coefficients come from columns 5 and 6 of Table V. As in Table IV, "unconcentrated" markets have a Herfindahl–Hirschmann Index (HHI) in the first quartile of the distribution (0.14); "concentrated" markets have an HHI in the top quartile (0.24). Regulated markets allow neither branching nor interstate banking; deregulated markets permit both.

	External Financial Dependence Measure Based on					
	Whether Industry Has Positive External Financial Dependence for Mature Compustat Firms		Level of External Financial Dependence for Mature Compustat Firms		Industry Median Loans/Assets for Firms in the 1998 SSBF	
	Number	Size	Number	Size	Number	Size
Relative effect (%) for a change from fully regulated to fully deregulated	11.6%	−12.3%	0.4%	−4.1%	1.2%	−22.0%
Relative effect (%) of a change from concentrated to unconcentrated banking market	4.6%	−9.3%	3.6%	−6.9%	2.3%	−7.3%

increased the number of establishments by 11.6% in the dependent sectors relative to the nondependent sectors, and reduced relative establishment size by 12.3%. Reducing banking concentration from 0.24 to 0.14 increases the number of firms by 4.6% and reduces average firm size by 9.3%.

The next four columns repeat this exercise using the two alternate measures of external financial dependence. Dependence equals the level of external funds needed to finance investment (columns 3 and 4) or the median loans/assets ratio for small firms represented in the 1998 SSBF (columns 5 and 6). Since these are continuous variables, we report the relative effects of increased banking competition comparing industries at the 75th percentile of the external financial dependence (loan/assets) distribution, with its effects on industries at the 25th percentile of the distribution. For example, we are comparing how bank competition affects a financially dependent industry like petroleum and coal products (SIC = 29) with one such as Food and Kindred Product (SIC = 20). As noted above in Table II, the level of external financial dependence is skewed by three "cash cow" industries that generate substantially more funds than they need (leather, tobacco, and apparel). Nevertheless, the economic significance of a decrease in banking concentration is only slightly smaller than what comes from the indicator variable models; the number of establishments increases by 3.6% with lower bank concentration and size declines by 6.9%. In contrast, the effects of deregulation are considerably smaller in these models. However, if we reduce the influence of the three outlier industries by setting their external financial dependence to −0.24 (the level of external financial dependence for the fourth-least dependent sector), the economic significance moves much closer to what comes out of the indicator variable specifications. Moving from full regulation to full deregulation now increases the number of establishments by 6.2% and decreases the average size by 8.2% (figures not reported in the table).

In the last two columns, economic significance is somewhat smaller for reductions in banking concentration than what comes from the indicator variable model. For deregulation, the effects are larger based on establishment size but smaller based on the number of establishments. It is worth repeating, however, that we view these results as robustness tests only because the use of external bank finance (loans) by small firms may reflect to some extent supply-side constraints stemming from the absence of banking competition.

C. Characterizing Changes in the Establishment Size Distribution

The results so far indicate that improvements in competitive conditions in banking lead to more establishments and a smaller average establishment size. These two changes could reflect entry by very small establishments; if so, then we would expect an increase in mass at the smallest end of the size distribution, and declines in mass elsewhere in the distribution. If better bank competition also helps the existing small firms grow (due to an increased supply of financial resources), then we ought to see a greater proportion not only of the smallest but also of midsized establishments as well. Moreover, testing for shifts in the whole size distribution allows us to compare how the shares of small- and mid-sized (presumably bank dependent) establishments behave relative to another sort of control group, namely, the share of the very largest establishments. These establishments (those with 1,000 or more employees) should not be affected by banking conditions because very large firms have access to nationwide (and competitive) securities markets. Thus, their fortunes should not vary with local credit conditions.

To explore these issues, we report the effects of changes in bank competition on the share of establishments at different points along the employment-size distribution (Tables VII and VIII). Each column in the two tables corresponds to the proportion of establishments in a given size bin. For example, in column 1 the dependent variable is based on the share of establishments with fewer than 5 employees, in column 2 the share of establishments with 5–19 employees,

Table VII

Regression of Logit of Size Shares on Bank Competition Variables

This table reports the coefficients from fixed effects regressions of the logit transformation of the share of establishments in different size bins in a state-industry-year. We do not report the logit for the share of establishments with more than 1,000 employees because about three quarters of the observations are zero. Data on establishments are from the County Business Patterns, 1977–1994. External financial dependence is an indicator for industries with above-median use of external finance as a fraction of capital expenditures (see Table II). The Herfindahl–Hirschmann Index (HHI), our measure of concentration, is built from HHIs at the city level (Metropolitan Statistical Area (MSA)). By city, we construct the sum of squared share of deposits for each banking company. We then average this concentration measure across all MSAs within the state, weighted by total deposits in each MSA (Source: FDIC Summary of Deposits.). The post-branching deregulation indicator equals one during all years in which a state permits in-state branching; the post-interstate banking deregulation indicator equals one during all years in which a state permits out-of-state banking companies to buy banks headquartered in the state. For a detailed discussion of the banking deregulation, see Kroszner and Strahan (1999). Industry share of employment equals the total employment in a given industry-state-year divided by total employment in the corresponding state-year. Absolute values of t-statistics are in parentheses, where * indicates significance at the 5% level and ** indicates significance at the 1% level.

	Share of Establishments with				
	Fewer than 5 Employees	5–19 Employees	20–99 Employees	100–999 Employees	1,000 or More Employees
Post-branching × external financial dependence indicator	0.063 (1.57)	−0.002 (0.06)	−0.053 (1.45)	−0.040 (0.94)	N/A
Post-interstate × external financial dependence indicator	−0.052 (0.85)	0.138 (2.54)**	0.111 (2.00)*	−0.271 (4.18)**	N/A
Local-market HHI × external financial dependence indicator	−0.929 (3.91)**	−0.492 (2.35)**	−0.575 (2.69)**	1.692 (6.76)**	N/A
Industry share of employment	−3.283 (19.36)**	−1.355 (9.05)**	2.608 (17.04)**	7.099 (39.71)**	N/A
Observations	15,127	15,127	15,127	15,127	
Fixed effects			State × Year Industry × Year		
R^2	0.21	0.25	0.21	0.33	

etc. To mitigate the possibility of heteroskedasticity, in Table VII we transform these proportions using the logit function (i.e., the $\log\{P/(1-P)\}$, where P equals the proportion of establishments in a given size bin). Note that because we disaggregate the data both by industry and state, there are a large number of observations with zero establishments in the over 1,000 category. Because the logit model is not feasible for this group, we also report the results based on the raw proportions (Table VIII). These coefficients are also easier to interpret than the logit coefficients.

The results in Table VII are again consistent with the earlier finding that increases in banking competition lead to increases in the importance of small

Table VIII

Regression of Size Shares on Bank Competition Variables

This table reports coefficients from fixed effects regressions of the share of establishments in different size bins in a state-industry-year. Data on establishments are from the County Business Patterns, 1977–1994. External financial dependence is an indicator for industries with above-median use of external finance as a fraction of capital expenditures (see Table II). The Herfindahl–Hirschmann Index (HHI), our measure of concentration, is built from HHIs at the city level (Metropolitan Statistical Area (MSA)). By city, we construct the sum of squared share of deposits for each banking company. We then average this concentration measure across all MSAs within the state, weighted by total deposits in each MSA (Source: FDIC Summary of Deposits). The post-branching deregulation indicator equals one during all years in which a state permits in-state branching; the post-interstate banking deregulation indicator equals one during all years in which a state permits out-of-state banking companies to buy banks headquartered in the state. For a detailed discussion of the banking deregulation, see Kroszner and Strahan (1999). Industry share of employment equals the total employment in a given industry-state-year divided by total employment in the corresponding state-year. Absolute values of t-statistics are in parentheses, where * indicates significance at the 5% level and ** indicates significance at the 1% level.

	Share of Establishments with				
	Fewer than 5 Employees	5–19 Employees	20–99 Employees	100–999 Employees	1,000 or More Employees
Post-branching × external financial dependence indicator	0.010 (2.03)*	−0.001 (0.11)	−0.007 (1.83)	−0.002 (0.49)	−0.001 (0.92)
Post-interstate × external financial dependence indicator	−0.012 (1.72)	0.019 (3.10)**	0.009 (1.71)	−0.018 (3.01)**	0.001 (0.99)
Local-market HHI × external financial dependence indicator	−0.083 (2.96)**	−0.015 (0.64)	−0.040 (1.84)	0.146 (6.44)**	−0.007 (1.59)
Industry share of employment	−0.609 (30.58)**	−0.323 (18.74)**	0.222 (14.46)**	0.610 (37.71)**	0.100 (32.98)**
Observations	15,127	15,127	15,127	15,127	15,127
Fixed effects			State × Year Industry × Year		
R^2	0.37	0.26	0.29	0.41	0.23

firms. That is, the size distribution of establishments shifts to the left (toward small establishments) as the banking industry becomes increasingly competitive. Reduced bank concentration increases the share of firms in the first three size bins, that is, fewer than 5 employees, 5–19 employees, and 20–99 employees (i.e., the sign of the coefficient on local-market HHI is negative). The effect of concentration is greatest for the smallest, and presumably most bank-dependent, size group. However, the increase in the share of firms in the next two size bins indicates that small incumbent firms also benefit from greater banking competition (assuming that most entry occurs among the smallest establishment). Conversely, we observe a decline in the share of firms in the 100–999 employee-size group with decreases in banking concentration (i.e., a positive coefficient in the HHI).

In contrast to concentration, bank deregulation, particularly interstate banking reform, seems to be positively related to the share of firms in the second- and third-smallest categories, whereas there is no significant effect on the smallest establishments. This increase in small establishments again occurs at the expense of firms with 100–999 employees. Interstate banking deregulation allows entry by large out-of-state banking organizations, which may explain why this reform seems to aid the medium-sized establishments more than the smallest establishments. Branching reform allows banks to expand, but only within state boundaries (at least until 1995, when interstate branching began). Thus, branching reform enhances competition but has much less of an impact on the presence of large banks. Its effects seem most concentrated on the smallest establishments. (Branching deregulation, while not significant in Table VII, is positive and significant for the smallest bin in Table VIII.)

Table VIII reports the results based on the raw share of establishment in each size bin. Here, we lose some statistical significance but gain a convenient way to assess the economic magnitudes of the results. For example, moving bank concentration from the 75th to the 25th percentile (a decline of 0.1) comes with an increase in the share of establishments in the smallest size bin of 0.83 percentage points, or a 3% increase in the unconditional mean (31%). The same decline in concentration comes with an increase of 0.15 percentage points in the share in the second smallest size bin and an increase of 0.40 percentage points in the third-smallest bin. The increased presence of small firms seems to come again from firms in the 100–999 employee group. According to Table VIII, this share declines by about 1.5 percentage points, which is about 12% of the unconditional mean (13%).

Table VIII also lets us test whether another control group—the largest establishments—is affected by changes in local banking competition. In fact, we find no significant change in the relative importance of the very largest establishments, those with more than 1,000 employees. This lack of reaction to changes in banking competition makes sense because the largest establishments are likely to be part of firms with access to nationwide securities markets. Hence, their access to financial resources does not depend on conditions in local banking markets.

As a final test, we consider whether regulatory changes alter the effects of banking concentration on the establishment size distribution. Black and Strahan (2002) argue that after states opened up their local banking markets to outside entry, the effects of concentration ought to have been mitigated by the threat of entry. That is, banking concentration no longer signals market power when barriers to entry from regulations have been eliminated. In fact, Black and Strahan find that the effect of concentration on the rate of creation of new incorporations does fall significantly with deregulation. We run similar tests using our measures of the size distribution and the number of establishments. Specifically, we add variables that interact the deregulation variables with our measure of bank concentration (and the industry external financial dependence measures). In these regressions (not reported), we estimate a generally positive coefficient (six of eight) on these interaction terms, suggesting that the effects of concentration on the size distribution may be attenuated by regulatory reform. These additional interaction effects, however, are not statistically significant.

V. Conclusions

We find that more vigorous competition in local U.S. banking markets—lower MSA-level bank concentration and looser state-level restrictions on bank entry—reduces the size of the typical establishment. Better bank competition also increases the share of establishments in the smallest size group and increases the total number of establishments. Significantly, changes in bank competition have no effect on the largest establishments, which makes sense given their access to financial resources in the commercial paper, corporate bond, and equity markets.

While theory does not paint a clear picture about how competition in banking ought to affect the firm-size distribution, the empirical work does. Comparing industry structure across local markets within the United States, or comparing structure across a large number of countries (both developed and developing), one reaches the same conclusion. Our empirical evidence is consistent with the idea that banks with market power erect an important financial barrier to entry to the detriment of the entrepreneurial sector of the economy, perhaps in part to protect the profitability of their existing borrowers. The evidence thus indicates that bank competition has a significant impact on important structural characteristics of sectors of production. Moreover, it indicates that such impact is not uniform across firms, but rather that depending on the degree of bank competition, some firms may benefit while others may lose. This is an important insight which updates the conventional wisdom that bank competition is either good or bad overall.

The policy implications associated with this issue are especially relevant. Banking market structure is a traditional policy variable whose control regulators across countries and over time often attempt to influence, sometimes in conflicting ways. For example, in the United States bank mergers have sometimes been altered to avoid excessive concentration in local markets. At the same time, however, until the 1980s many states protected their banks from competition through branching and interstate banking restrictions. Similar restraints on competition have been common elsewhere; for example, many countries continue to protect their banks from foreign entry. One can well understand why political forces lead to tight restraints on banking competition if both incumbent banks and incumbent firms benefit from the restraints. In fact, Rajan and Zingales (2004) use historical evidence to argue broadly that incumbent firms often fight hardest to prevent financial openness, sometimes leading to long-term declines in a country's growth prospects.

REFERENCES

Beck, Thorsten, Asli Demirguc-Kunt, and Vojislav Maksimovic, 2004, Bank competition and access to finance: International evidence, *Journal of Money, Credit, and Banking* 36, 627–648.

Berger, Allen N., Rebecca S. Demsetz, and Philip E. Strahan, 1999, The consolidation of the financial services industry: Cause, consequences and implications for the future, *Journal of Banking and Finance* 23, 135–194.

Bertrand, Marianne, Antoinette Schoar, and David Thesmar, 2004, Banking deregulation and industry structure: Evidence from the French banking reforms of 1985, C.E.P.R. Discussion Papers, 4488.

Bhattacharya, Sudipto, and Gabriella Chiesa, 1995, Proprietary information, financial intermediation, and research incentives, *Journal of Financial Intermediation* 4, 328–357.

Bitler, Marianne P., Alicia M. Robb, and John D. Wolken, 2001, Financial services used by small businesses: Evidence from the 1998 survey of small business finance, *Federal Reserve Bulletin* 87, 183–205.

Black, Sandra E., and Philip E. Strahan, 2002, Entrepreneurship and bank credit availability, *Journal of Finance* 57, 2807–2833.

Bonaccorsi di Patti, Emilia, and Giovanni Dell'Ariccia, 2004, Bank competition and firm creation, *Journal of Money, Credit, and Banking* 36, 225–251.

Brander, James, and Tracy Lewis, 1986, Oligopoly and financial structure: The limited liability effect, *The American Economic Review* 76, 956–970.

Campbell, Jeffrey R., and Hugo A. Hopenhayn, 2003, Market size matters, Federal Reserve of Chicago, Working paper no. 2003-12.

Cestone, Giacinta, and Lucy White, 2003, Anti-competitive financial contracting: The design of financial claims, *Journal of Finance* 58, 2109–2142.

Cetorelli, Nicola, 2001, Does bank concentration lead to concentration in industrial sectors? Working paper 01-01, Federal Reserve Bank of Chicago.

Cetorelli, Nicola, 2003, Life-cycle dynamics in industrial sectors: The role of banking market structure, *Quarterly Review, Federal Reserve Bank of St. Louis* 85, 135–147.

Cetorelli, Nicola, 2004, Real effects of bank competition, *Journal of Money, Credit, and Banking* 36, 543–558.

Cetorelli, Nicola, and Michele Gambera, 2001, Banking market structure, financial dependence and growth: International evidence from industry data, *Journal of Finance* 56, 617–648.

Chevalier, Judith, 1995, Capital structure and product market competition: Empirical evidence from the supermarket industry, *American Economic Review* 85, 415–435.

Demirguc-Kunt, Asli, and Vojislav Maksimovic, 1998, Law, finance and firm growth, *Journal of Finance* 53, 2107–2137.

Evans, David S., 1987, The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries, *Journal of Industrial Economics* 35, 567–581.

Goldsmith, Raymond, 1969, *Financial Structure and Development* (Yale University Press, New Haven).

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2004a, The role of social capital on financial development, *American Economic Review* 94, 526–556.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2004b, Does local financial development matter? *Quarterly Journal of Economics* 119, 929–969.

Gurley, John G., and Edward S. Shaw, 1955, Financial aspects of economic development, *American Economic Review* 45, 515–538.

Hellmann, Thomas, and Marco Da Rin, 2002, Banks as catalysts for industrialization, *Journal of Financial Intermediation* 11, 366–397.

Hopenhayn, Hugo A., 1992, Entry, exit, and firm dynamics in long run equilibrium, *Econometrica* 60, 1127–1150.

Jayaratne, Jith, and Philip E. Strahan, 1996, The finance-growth nexus: Evidence from bank branch deregulation, *Quarterly Journal of Economics* 101, 639–670.

Jayaratne, Jith, and Philip E. Strahan, 1998, Entry restrictions, industry evolution and dynamic efficiency: Evidence from commercial banking, *Journal of Law and Economics* 41, 239–274.

Jovanovic, Boyan, 1982, Selection and evolution of industry, *Econometrica* 50, 649–670.

King, Robert, and Ross Levine, 1993a, Finance, entrepreneurship and growth, *Journal of Monetary Economics* 32, 513–542.

King, Robert, and Ross Levine, 1993b, Finance and growth: Schumpeter might be right, *Quarterly Journal of Economics* 108, 717–738.

Klapper, Leora, Luc Laeven, and Raghuram G. Rajan, 2004, Barriers to entrepreneurship, mimeo, November 2004.

Klepper, Steven, 1996, Entry, exit, growth, and innovation over the product life cycle, *American Economic Review* 86, 562–583.

Kovenock, Dan, and Gordon Phillips, 1995, Capital structure and product market rivalry: How do we reconcile theory and evidence? *American Economic Review* 85, 403–408.

Kovenock, Dan, and Gordon Phillips, 1997, Capital structure and product market behavior: An examination of plant closing and investment decisions, *Review of Financial Studies* 10, 767–803.

Kroszner, Randall S., and Philip E. Strahan, 1999, What drives deregulation? Economics and politics of the relaxation of bank branching restrictions, *Quarterly Journal of Economics* 114, 1437–1467.

Kumar, Krishna, Raghuram G. Rajan, and Luigi Zingales, 2001, What determines firm size? NBER Working paper no. 7208.

Laporta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2002, Government ownership of banks, *Journal of Finance* 57, 265–301.

Levine, Ross, 1999, Law, finance and economic growth, *Journal of Financial Intermediation* 8, 8–35

Levine, Ross, Norman Loayza, and Thorsten Beck, 2000, Financial intermediation and growth: Causality and causes, *Journal of Monetary Economics* 46, 31–77.

Levine, Ross, and Sara Zervos, 1998, Stock markets, banks and economic growth, *American Economic Review* 88, 537–558.

Lucas, Robert E., 1978, On the size distribution of firms, *Bell Journal of Economics* 9, 508–523.

Maksimovic, Vojislav, 1988, Capital structure in a repeated oligopoly, *Rand Journal of Economics* 3, 389–407.

Maurer, Noel, and Stephen Haber, 2003, Bank concentration, related lending, and economic performance: Evidence from Mexico, mimeo, Stanford University.

Mayer, Colin, 1988, New issues in corporate finance, *European Economic Review* 32, 1167–1183.

McKinnon, Ronald, 1973, *Money and Capital in Economic Development* (Brookings Institution, Washington, DC).

Peltzman, Sam, 1969, The banking structure and the transmission of monetary policy, *Journal of Finance* 24, 387–411.

Petersen, Mitchell A., and Raghuram G. Rajan, 1995, The effect of credit market competition on lending relationships, *Quarterly Journal of Economics* 110, 407–443.

Petersen, Mitchell A., and Raghuram G. Rajan, 2002, Does distance still matter? The information revolution in small business lending, *Journal of Finance* 57, 2533–2570.

Rajan, Raghuram G., and Luigi Zingales, 1998, Financial dependence and growth, *American Economic Review* 88, 559–587.

Rajan, Raghuram G., and Luigi Zingales, 2004, *Saving Capitalism from the Capitalists*: Unleashing the Power of Financial Markets to Create Wealth and Spread Opportunity (Princeton University Press, Princeton, NJ).

Schumpeter, Joseph A., 1912, *Theorie der Wirtschaftlichen Entwicklung. Leipzig: Dunker & Humblot* (The Theory of Economic Development, translated by Redvers Opie, 1934. Harvard University Press, Cambridge, MA).

Spagnolo, Giancarlo, 2000, Debt as a (credible) collusive device, Working paper, 243, Stockholm School of Econommics.

Stiroh, Kevin, and Philip E. Strahan, 2003, Competitive dynamics of deregulation: Evidence from U.S. banking, *Journal of Money, Credit, and Banking* 35, 801–828.

Zombie Lending and Depressed Restructuring in Japan

Ricardo J. Caballero

Massachusetts Institute of Technology and NBER

Takeo Hoshi

University of California at San Diego, Graduate School of International Relations and Pacific Studies and NBER

Anil K Kashyap

University of Chicago, Graduate School of Business, Federal Reserve Bank of Chicago and NBER

This draft: March 2006

We thank numerous seminar participants and our colleagues for useful comments. We especially benefited from helpful comments by Olivier Blanchard, Roger Bohn, Toni Braun, Kenichiro Kobayashi, Hugh Patrick, and Makoto Sakuragawa. We thank Yoichi Arai, Munechika Katayama and Tatsuyoshi Okimoto for expert research assistance. Caballero thanks the National Science Foundation for research support. Hoshi thanks the Research Institute of Economy, Trade, and Industry (RIETI) for research support. Kashyap thanks the Center for Research in Securities Prices and the Stigler Center both at the University of Chicago Graduate School of Business for research support. This research was also funded in part by the Ewing Marion Kauffman Foundation. The views expressed in this paper are those of the authors and not necessarily of any of the organizations with which we are affiliated or which sponsored this research. Future drafts of this paper will be posted to http://gsbwww.uchicago.edu/fac/anil.kashyap/research. First draft: September 2003.

Zombie Lending and Depressed Restructuring in Japan

Abstract:

In this paper, we propose a bank-based explanation for the decade-long Japanese slowdown following the asset price collapse in the early 1990s. We start with the well-known observation that most large Japanese banks were only able to comply with capital standards because regulators were lax in their inspections. To facilitate this forbearance the banks often engaged in sham loan restructurings that kept credit flowing to otherwise insolvent borrowers (that we call zombies). Thus, the normal competitive outcome whereby the zombies would shed workers and lose market share was thwarted. Our model highlights the restructuring implications of the zombie problem. The counterpart of the congestion created by the zombies is a reduction of the profits for healthy firms, which discourages their entry and investment. In this context, even solvent banks will not find good lending opportunities. We confirm our story's key predictions that zombie-dominated industries exhibit more depressed job creation and destruction, and lower productivity. We present firm-level regressions showing that the increase in zombies depressed the investment and employment growth of non-zombies and widened the productivity gap between zombies and non-zombies.

1. Introduction

This paper explores the role that misdirected bank lending played in prolonging the Japanese macroeconomic stagnation that began in the early 1990s. The investigation focuses on the widespread practice of Japanese banks of continuing to lend to otherwise insolvent firms. We document the prevalence of this forbearance lending and show its distorting effects on healthy firms that were competing with the impaired firms.

Hoshi (2000) was the first paper to call attention to this phenomenon and its ramifications have been partially explored by a number of observers of the Japanese economy. There is agreement that the trigger was the large stock and land price declines that began in early 1990s: stock prices lost roughly 60% of their value from the 1989 peak within three years, while commercial land prices fell by roughly 50% after their 1992 peak over the next ten years. These shocks impaired collateral values sufficiently that any banking system would have had tremendous problems adjusting. But in Japan the political and regulatory response was to deny the existence of any problems and delay any serious reforms or restructuring of the banks.[1] Aside from a couple of crisis periods when regulators were forced to recognize a few insolvencies and temporarily nationalize the offending banks, the banks were surprisingly unconstrained by the regulators.

The one exception to this rule is that banks had to comply (or appear to comply) with the international standards governing their minimum level of capital (the so-called Basle capital standards). This meant that when banks wanted to call in a non-performing loan, they were likely to have to write off existing capital, which in turn pushed them up against the minimum capital levels. The fear of falling below the capital standards led many banks to continue to extend credit to insolvent borrowers, gambling that somehow these firms would recover or that the government would bail them out.[2] Failing to

[1] For instance, in 1997, at least 5 years after the problem of non-performing loans was recognized, the Ministry of Finance was insisting that no public money would be needed to assist the banks. In February 1999 then Vice Minister of International Finance, Eisuke Sakakibara, was quoted as saying that the Japanese banking problems "would be over within a matter of weeks." As late as 2002, the Financial Services Agency claimed that Japanese banks were well capitalized and no *more* public money would be necessary.

[2] The banks also tried to raise capital by issuing more shares and subordinated debt, as Ito and Sasaki (2002) document. When the banks raised new capital, however, almost all came from either related firms (most notably life insurance companies) that are dependent on the banks for their financing, or the

rollover the loans also would have sparked public criticism that banks were worsening the recession by denying credit to needy corporations. Indeed, the government also encouraged the banks to *increase* their lending to small and medium sized firms to ease the apparent "credit crunch" especially after 1998.[3] The continued financing, or "ever-greening," can therefore be seen as a rational response by the banks to these various pressures.

A simple measure of the ever-greening is shown in Figure 1, which reports the percentage of bank customers that received subsidized bank credit. We defer the details of how the firms are identified until the next section, but for now all that matters is that the universe of firms considered here is all publicly traded manufacturing, construction, real estate, retail, wholesale (excluding nine general trading companies) and service sector firms. The top panel of the figure shows roughly 30% of these firms were on life support from the banks in the early 2000s. The lower panel, which shows comparable asset weighted figures, suggests that about 15% of assets reside in these firms. As these figures show, these percentages were much lower in the 1980s and early 1990s.

By keeping these unprofitable borrowers (that we call "zombies") alive, the banks allowed them to distort competition throughout the rest of the economy. The zombies' distortions came in many ways, including depressing market prices for their products, raising market wages by hanging on to the workers whose productivity at the current firms declined and, more generally, congesting the markets where they participated. Effectively the growing government liability that came from guaranteeing the deposits of banks that supported the zombies served as a very inefficient program to sustain employment. Thus, the normal competitive outcome whereby the zombies would shed workers and lose market share was thwarted.[4] More importantly, the low prices and high wages reduce the profits that new and more productive firms could earn, thereby

---

government when banks received capital injections. See Hoshi and Kashyap (2004, 2005) for more on this "double-gearing" between banking and life insurance sectors.

[3] Subsequently when the Long-Term Credit Bank was returned to private ownership, a condition for the sale was the new owners would maintain lending to small and medium borrowers. The new owners tightened credit standards and the government pressured them to continue supplying funds, see Tett (2003) for details.

[4] See Ahearne and Shinada (2004) for some direct evidence suggesting that inefficient firms in the non-manufacturing sector gained market share in Japan in the 1990s. See also Kim (2004) and Restuccia and Rogerson (2003) for attempts to quantify the size of these types of distortions.

discouraging their entry and investment. In addition, even solvent banks saw no particularly good lending opportunities in Japan.

In the remainder of the paper we document and formalize this story. In the next section, we describe the construction of our zombie measure. There are a number of potential proxies that could be used to identify zombies. As we explain, however, measurement problems confound most of these alternatives.

Having measured the extent of zombies, we then model their effects. The model is a standard variant of the type that is studied in the literature on creative destruction. It is designed to contrast the adjustment of an industry to a negative shock with and without the presence of zombies. We model the presence of zombies as a constraint on the natural surge in destruction that would arise in the wake of an unfavorable technological, demand, or credit shock. The main effect of that constraint is that job creation must slow sufficiently to re-equilibrate the economy. This means that during the adjustment the economy is characterized by what Caballero and Hammour (1998, 2000) have called "sclerosis" — the preservation of production units that would not be saved without the banks' subsidies— and the associated "scrambling" — the retention of firms and projects that are less productive than some of those that do not enter or are not implemented due to the congestion caused by the zombies.

In the fourth section of the paper, we assess the main aggregate implications of the model. In particular, we study the interaction between the percentage of zombies in the economy and the amount of restructuring, both over time and across different sectors. We find that the rise of the zombies has been associated with falling levels of aggregate restructuring, with job creation being especially depressed in the parts of the economy with the most zombies firms. We then explore the impact of zombies on sectoral performance measures. We find that the prevalence of zombies lowers productivity.

In section 5 we analyze firm-level data to directly look for congestion effects of the zombies on non-zombie firms' behavior. We find that investment and employment growth for healthy firms falls as the percentage of zombies in their industry rises. Moreover, the gap in productivity between zombie and non-zombie firms rises as the percentage of zombies rises. Most strikingly, the presence of the zombies depresses activity the most for the fastest growing healthy firms. All of these findings are consistent with the predictions that zombies crowd the market and that the congestion has real effects on the healthy firms in the economy. Simple extrapolations using our regression coefficients suggest that cumulative size of the distortions (in terms of investment, or employment) is substantial.

In the final section of the paper we conclude by summarizing our results and discussing the implications of our findings for Japan's outlook.

2. Identifying zombies

Our story can be divided into two parts. First, the banks misallocated credit by supporting zombie firms. Second, the existence of zombie firms interfered with the process of creative destruction and stifled growth. Our measure of zombie should not only capture the misallocation of credit but also be useful in testing the effect of zombies on corporate profitability and growth.

2.1 Defining Zombies

There is a growing literature examining the potential misallocation of bank credit in Japan (see Sekine, Kobayashi, and Saita (2003) for a survey). Much of the evidence is indirect. For instance, several papers (including Hoshi (2000), Fukao (2000), Hosono and Sakuragawa (2003), Sasaki (2004)) study the distribution of loans across industries and note that underperforming industries like real estate or construction received more bank credit than other sectors that were performing better (such as manufacturing).[5]

Peek and Rosengren (2005) offer the most direct and systematic study to date on the potential misallocation of bank credit. They find that bank credit to poor performing firms often increased between 1993 and 1999. These firms' main banks are more likely to lend to the firms than other banks dealing with these firms when the firm's profitability

[5] Other indirect evidence comes from studies such as Smith (2003), Schaede (2005) and Jerram (2004) that document that loan rates in Japan do not appear to be high enough to reflect the riskiness of the loans. Sakai, Uesugi and Watanabe (2005), however, show that poorly performing firms (measured by operating profits or net worth) still pay higher bank loan rates and are more likely to exit compared with better performing firms, at least for small firms. Finally, see also Hamao, Mei and Xu (forthcoming) who show that firm-level equity returns became less volatile during the 1990s and argue that this is likely due to a lack of restructuring in the economy.

is declining. This pattern of perverse credit allocation is more likely when the bank's own balance sheet is weak or when the borrower is a keiretsu affiliate. Importantly, non-affiliated banks do not show this pattern.

We depart from past studies by trying to identify zombies by classifying firms only based on our assessment of whether they are receiving subsidized credit, and not by looking at their productivity or profitability. This strategy permits us to evaluate the effect of zombies on the economy. If instead we were to define zombies based on their operating characteristics, then almost by definition industries dominated by zombie firms would have low profitability, and likely also have low growth. Rather than hard-wiring this correlation, we want to test for it.

The challenge for our approach is to use publicly available information to determine which firms are receiving subsidized credit: banks and their borrowers have little incentive to reveal that a loan is miss-priced. Because of the myriad of ways in which banks could transfer resources to their clients, there are many ways that we could attempt to measure subsidies. To get some guidance we used the Nikkei Telecom 21 to search the four newspapers published by the Nihon Keizai Shimbun-sha (*Nihon Keizai Shimbun, Nikkei Kin'yū Shimbun, Nikkei Sangyō Shimbun, Nikkei Ryūtsū Shimbun*) between January 1990 and May 2004 for all news articles containing the words "financial assistance" and either "management reconstruction plan" or ("corporation" and "reconstruction").[6] The summary of our findings are given in Table 1.

Our search uncovers 120 separate cases. In most of them there were multiple types of assistance that were included. As the table shows, between interest rate concessions, debt-equity swaps, debt forgiveness, and moratoriums on loan principal or interest, most of these packages involve reductions in interest payments or outright debt forgiveness by the troubled firms.[7]

The decision by a bank to restructure the loans to distressed companies in these ways, rather than just rolling over the loans, helps reduce the required capital needed by the bank. Without such restructuring, banks would be forced to classify the loans to

[6] The Japanese phrases were Kin'yu Shien AND (Keiei Saiken Keikaku OR (Kigyo AND Saiken)).

[7] These patterns are consistent with the claim by Tett and Ibison (2001) that almost one-half of the public funds injected into the banking system in 1998 and 1999 were allowed to be passed on to troubled construction companies in the form of debt forgiveness.

those borrowers as “at risk”, which usually would require the banks to set aside 70% of the loan value as loan loss reserves. With restructuring, the banks need only move the loans to the “special attention” category, which requires reserves of at most 15%.

In light of the evidence in Table 1, we concentrate on credit assistance that involves a direct interest rate subsidy. We proceed in three steps. First, we calculate a hypothetical lower bound for interest payments (R^*) that we expect only for the highest quality borrowers. We then compare this lower bound to the observed interest payments. Finally, we make several econometric assumptions to use the observed difference between actual interest rate (r) and notional lower bound rate (r^*) to infer cases where we believe subsidies are present.

2.2 Detecting Zombies

The minimum required interest payment for each firm each year, $R*_{i,t,}$ is defined as:

$$R*_{i,t} = rs_{t-1}BS_{i,t-1} + \left(\tfrac{1}{5}\sum_{j=1}^{5} rl_{t-j} \right) \cdot BL_{i,t-1} + rcb_{\text{min over last 5 years},t} * Bonds_{i,t-1}$$

where $BS_{i,t}$, $BL_{i,t}$ and $Bonds_{i,t}$ are short-term bank loans (less than one year), long-term bank loans (more than one year), and total bonds outstanding (including convertible bonds (CBs) and warrant-attached bonds) respectively of firm i at the end of year t, and rs_t , rl_t , and $rcb_{\text{min over the last 5 years, } t}$ are the average short-term prime rate in year t, the average long-term prime rate in year t, and the minimum observed coupon rate on any convertible corporate bond issued in the last five years before t.

This estimate for the lower bound reflects the data constraints we face. In particular, all we know about the firms’ debt structure is the type of debt instrument (short-term bank borrowing, long-term borrowing that are due in one year and remaining long-term bank borrowing, bonds outstanding that are due in one year and remaining bonds outstanding, and commercial paper outstanding). In other words, we do not know the exact interest rates on specific loans, bonds or commercial paper, nor do we know the exact maturities of any of these obligations. Finally, the interest payments we can measure include all interest, fee and discount expenses, including those related to trade credit.

The general principle guiding the choices we make is to select interest rates that are extremely advantageous for the borrower, so that R^* is in fact less than what most firms would pay in the absence of subsidies. For instance, by assuming that bond financing takes place at $rcb_{\text{min over the last 5 years, } t}$ we are assuming not only that firms borrow using convertible bonds (which carry lower interest rates due to the conversion option), but also that these bonds are issued when rates are at their lowest. We provide additional discussion of the data choices used in constructing R^* and the alternative approaches that we examined for robustness check in Appendix 1.

To categorize firms we compare the actual interest payments made by the firms ($R_{i,t}$) with our hypothetical lower bound. We normalize the difference by the amount of total borrowing at the beginning of the period ($B_{i.t-1} = BS_{i,t-1} + BL_{i,t-1} + Bonds_{i,t-1} + CP_{i,t-1}$), where $CP_{i,t-1}$ is the amount of commercial paper outstanding for the firm i at the beginning of the period t, so that the units are comparable to interest rates. Accordingly we refer to the resulting variable, $x_{i,t} \equiv \frac{R_{i,t} - R^*_{i,t}}{B_{i,t-1}} = r_{i,t} - r^*_{i,t}$, as the interest rate gap. This measure is "conservative" because we assume the minimum interest rates that are extremely advantageous to the firm and because the interest payment, $R_{i,t}$, includes interest expenses on items beyond our concept of total borrowing (such as interest expenses on trade credit).

Note that given our procedure to construct r^* we will not be able to detect all types of subsidized lending.[8] In particular, any type of assistance that lowers the current period's interest payments can be detected: including debt forgiveness, interest rate concessions, debt for equity swaps, or moratoriums on interest rate payments, all of which appeared to be prevalent in the cases studied in Table 1. On the other hand, if a

---

[8] In addition to the cases studied below, Hoshi (2006) examines the potential problems that might arise from rapid changes in interest rates. For example, if interest rates fell sharply and actual loan terms moved as well, then our gap variable could be misleading about the prevalence of subsidized loans. He constructs an alternative measure (that would be more robust to within year interest rate changes) and concludes that this sort of problem does not appear to be quantitatively important.

bank makes new loans to a firm at normal interest rates that are then used to pay off past loans, then our gap variable will not capture the subsidy. Likewise, if a bank buys other assets from a client at overly generous prices our proxy will not detect the assistance.

We explore two strategies for identifying the set of zombie firms from the calculated interest rate gaps. Our baseline procedure classifies a firm i as a zombie for year t whenever its interest rate gap is negative ($x_{it} < 0$). The justification for this strategy is the conservative philosophy underlying the construction of r^*. If r^* is a perfectly measured lower bound, then only a firm that receives a subsidy can have a negative gap. However, the problem of labeling a firm with x_{it} just above zero as non-zombie remains even under this perfect scenario.

Thus we resort to a second approach, which is more robust to misclassification of non-zombies. In this second approach we assume that the set of zombies is a "fuzzy" set. In the classical set theory, an element either belongs or does not belong to a particular set so that a 0-1 indicator function can be used to define a subset. In contrast, in fuzzy set theory an element can belong to a particular subset to a certain degree, so that the indicator function can take any value in the interval [0, 1]. When the images of the indicator function are confined to {0, 1}, a set defined by the indicator function is called a "crisp" set. Using this terminology, our first approach assumes the set of zombies is "crisp." Our second approach, on the other hand, assumes the set is "fuzzy," allowing some firms to be more-or-less zombie-like.[9]

The indicator function that defines a fuzzy subset is called "membership function," which we assume to be (for the set of zombie firms):

$$z(x;d_1,d_2)=\begin{cases}1 & \text{if } x<d_1\\ \dfrac{d_2-x}{d_2-d_1} & \text{if } d_1\le x\le d_2 \quad \text{where } d_1\le 0\le d_2\\ 0 & \text{if } x>d_2\end{cases} \tag{1}$$

The shape of the membership function is determined by the two parameters, d_1 and d_2. Figure 2 shows this membership function along with the indicator function implicit in our first approach. It is easy to see the second approach degenerates to our first approach when d_1 and d_2 are both zero.

[9] See Nguyen and Walker (2006) for an introduction to the fuzzy set theory.

The second approach is appealing given the fuzzy nature of the concept of "zombie firms." These are defined to be those firms that receive sufficient financial help from their creditors to survive in spite of their poor profitability. It is inherently difficult to specify how much financial help is considered to be sufficient, even if we had access to much more information than we do about individual firms. Our fuzzy approach acknowledges this limitation and assigns numbers between 0 and 1 to those firms whose zombie status is ambiguous.

Given the asymmetry (toward conservatism) inherent in the construction of r^*, we assume that d_1 is closer to zero than d_2. In what follows we show results for $(d_1, d_2) = (0,$ 50bp$)$ and $(d_1, d_2) = ($-25bp, 75bp$)$, where bp stands for basis points. Thus, in the first case, we assume a firm with x_{it} below zero is a definite zombie and a firm with x_{it} above 50 basis points is definitely a non-zombie: any firm with x_{it} between zero and 50 basis points has "zombiness" between 0 and 1.

2.3 Quantifying the prevalence of zombies

Figure 1 shows the aggregate estimate of the percentage of zombies using our baseline procedure. As mentioned earlier, treating all firms equally we see that the percentage of zombies hovered between 5 and 15 percent up until 1993 and then rose sharply over the mid 1990s so that the zombie percentage was above 25 percent for every year after 1994. In terms of the congestion spillovers, a size weighted measure of zombies is likely to be more important. Weighting firms by their assets we see the same general pattern but with the overall percentage being lower, closer to 15 percent in the latter part of the sample.

We view the cross-sectional prevalence of zombies as another way to assess the plausibility of our definition. To conduct this assessment, we aggregated the data used in Figure 1 into five industry groups covering manufacturing, construction, real estate, retail and wholesale (other than the nine largest general trading companies), and services – recall that all the firms included here are publicly traded. The zombie index for an industry is constructed by calculating the share of total assets held by the zombie firms –

and for the remainder of the paper we concentrate on asset weighted zombie indices. In addition to showing the industry distribution, we also compute the zombie percentages implied by our second procedure with (d_1, d_2) = (0, 50bp) and (d_1, d_2) = (-25bp, 75bp).

Figure 3 shows the zombie index for each industry from 1981 to 2002. We draw three main conclusions from these graphs. Starting with the upper left hand panel that shows the data for the entire sample, first notice that the crisp zombie measure (our baseline case) and the two fuzzy measures share similar time series movements (with the correlation between the crisp measure and the two fuzzy measures exceeding 0.99). Second, the other five panels show that the proportion of zombie firms increased in the late 1990s in every industry. The third key conclusion is that the zombie problem was more serious for non-manufacturing firms than for manufacturing firms. In manufacturing, the crisp measure suggests that zombie index only rose from 3.11% (1981-1993 average) to 9.58% (1996-2002 average). In the construction industry, however, the measure increased from 4.47% (1981-1993 average) to 20.35% (1996-2002 average). Similar large increases occurred for the wholesale and retail, services, and real estate industries.

There are a variety of potential explanations for these cross-sectional differences. For instance, Japanese manufacturing firms face global competition and thus could not easily be protected without prohibitively large subsidies. For example, many of the troubled Japanese automakers were taken over by foreign firms rather than rescued by their banks during the 1990s. In contrast, there is very little foreign competition in the other four industries.

A second important factor was the nature of the shocks hitting the different sectors. For instance, the construction and real estate industries were forced to deal with the huge run-up and subsequent collapse of land prices mentioned earlier. Thus, the adjustment for these industries was likely to be more wrenching than for the other sectors.

But the most important point about the differences shown in Figure 3 is that they confirm the conventional wisdom that bank lending distortions were not equal across sectors and that the problems were less acute in manufacturing – see Sekine et al (2003) for further discussion. Thus, regardless of which explanation one favors as to why this might be the case, we view it as particularly reassuring that our zombie index confirms this conventional view.

Figure 4, our last plausibility check, shows the asset weighted percentages of zombies for the firms that are above and below the median profit rate for their industry. To keep the graphs readable we show only the crisp measures, but the other measures show similar patterns. In manufacturing the differences are not very noticeable, with slightly fewer high profit firms being labeled as zombies. In the remaining industries, particularly in real estate and construction, it appears that our measure of zombies is identifying firms that are systematically less profitable than the non-zombies, particularly from the mid-1990s onward.

3. A model of the effect of zombie firms on restructuring

To analyze the effect of zombies we study a very simple environment that involves entry and exit decisions of both incumbent firms and potential new firms, which we later extend to analyze expansion and contraction decisions of existing firms. As a benchmark we start with a normal environment where all decisions are based purely on the operating profits from running a firm. We then contrast that environment to one where some incumbent firms (for an unspecified reason) receive a subsidy that allows them to remain in business despite negative operating profits.

3.1 The Environment

The essential points of interest can be seen in a model where time is discrete (and indexed by t). A (representative) period t starts with a mass m_t of existing production units. The productivity of the incumbents varies over time and the current level of productivity for firm *i* in year *t*, y_{it}^{o}, is:

$$y_{it}^{o} = A + \epsilon_{it}^{o}$$

where $\in_{it}^{o}$ is an idiosyncratic shock that is distributed uniformly on the unit interval. The main predictions from this model do not depend on the persistence of the productivity shocks, so we assume these shocks are i.i.d.

In addition to the incumbents, there are also a set of potential entrants and we normalize their mass to be ½. The potential entrants each draw a productivity level, y_{it}^{n}, before deciding whether to enter or not. The productivity for i^th potential new firm in year *t* is:

$$y_{it}^{n} = A + B + \in_{it}^{n}$$

with $B > 0$ and $\in_{it}^{n}$ distributed uniformly on the unit interval. The shock $\in_{it}^{n}$ is again assumed to have no persistence. These assumptions imply that on average the potential new firms will be more productive (and more profitable) than the incumbents (for one period only, then they become incumbents as well). However, we also assume that there is an entry cost, $\kappa > 0$, that they must pay to start up.

Finally, both new and old units must incur a cost $p(N_t)$ in order to produce, where N_t represents the number of production units in operation at time t, i.e., the sum of the existing units that do not exit and new entrants. The cost $p(N)$ is increasing with respect to *N* and captures any scarce input such as land, labor or capital. Indeed *p(N)* captures any reduction in profits due to congestion or competition.[10] For our purposes, all the predictions we emphasize will hold as long as $p(N)$ is a strictly increasing continuous function of *N*. For simplicity, we adopt the linear function:

$$p(N_t) = N_t + \mu.$$

[10] For example, we can motivate *p(N)* as the reduction in profits due to competition in the output market. Suppose the price of output is given by $D^{-1}(N)$, a decreasing function of *N*, and that the cost of production for each production unit is a constant, *C*. Under our assumption on productivity, an incumbent decides to stay in the market (and a potential entrant decides to enter the market) if $D^{-1}(N)(A+\varepsilon) - C > 0$, or equivalently, $A+\varepsilon - C/D^{-1}(N) > 0$. In this specific example, *p(N)* is $C/D^{-1}(N)$, which is increasing with respect to *N*.

where the intercept μ is potential shift variable that captures cost changes and other profit shocks.

3.2 Decisions

This basic model will quickly generate complicated dynamics because the existing firms have paid the entry cost and thus face a different decision problem than the new firms for which the entry cost is not sunk. These dynamics are not essential for our main predictions, so we assume that $B = \kappa$. In this case, the exit decision by incumbents and the entry decision by potential entrants become fully myopic. Since productivity shocks are i.i.d. and there is no advantage from being an insider (the sunk cost of investment is exactly offset by a lower productivity), both types of units look only at current profits to decide whether to operate.

Letting $\overline{y}^o$ and $\overline{y}^n$ denote the reservation productivity of incumbents and potential entrants, respectively, we have:

$$\overline{y}^o - p(N) = 0,$$

$$\overline{y}^n - \kappa - p(N) = 0.$$

In this case it is straightforward to find the mass of exit, D_t, and entry, H_t, respectively:

$$D_t = m_t\left[1 - \int_{p(N_t)-A}^{1} di\right] = m_t(p(N_t) - A), \tag{2}$$

$$H_t = \frac{1}{2}\int_{p(N_t)-A}^{1} di = \frac{1}{2}(1 - (p(N_t) - A)). \tag{3}$$

Adding units created to the surviving incumbents yields the total number of units operating at time t:

$$N_t = H_t + m_t - D_t = \left(\frac{1}{2} + m_t\right)\left(1 - \left(p(N_t) - A\right)\right). \qquad (4)$$

3.3 Equilibrium and Steady State

We can now solve for the steady state of the normal version of the economy. The first step is to replace $p(N)$ with $N + \mu$ in (4). The notation is simplified if we define S to be composite shock that is equal to A-μ. Note that a lower S indicates either higher costs (higher μ) or lower average productivity (smaller A). This yields the equilibrium number of units:

$$N_t = \left(\frac{1/2 + m_t}{3/2 + m_t}\right)(1 + S). \qquad (5)$$

Given the total number of operating units, we can solve for equilibrium rates of destruction and creation by substituting (5) into (2) and (3):

$$D_t = m_t\left(\frac{1/2 + m_t - S}{3/2 + m_t}\right) \qquad (6)$$

$$H_t = \frac{1}{2}\left(\frac{1 + S}{3/2 + m_t}\right). \qquad (7)$$

The dynamics of this system are determined by:

$$m_{t+1} = N_t. \qquad (8)$$

In steady state, the mass of incumbents remains constant at $m^{ss} = N^{ss}$, which requires that creation and destruction exactly offset each other or, equivalently, that

$m_t = N_t$. Using the latter condition and (5), yields a quadratic equation for m^{ss}, which has a unique positive solution of:

$$m^{ss} = \frac{S - \frac{1}{2} + \sqrt{\left(\frac{1}{2} - S\right)^2 + 2(1+S)}}{2}$$

For small values of *S*, we can approximate the above by:

$$m^{ss} \approx \frac{1}{2} + \frac{2}{3}S.$$

In our subsequent analysis we will assume that the economy begins in a steady state and that the initial (pre-shock) value of S, S_0, is 0. Given this normalization, the corresponding steady state will be $m_0 = N_0 = 1/2$ and $H_0 = D_0 = 1/4$.

3.4 A (permanent) Recession

We can now analyze the adjustment of the economy to a profit shock. By construction the model treats aggregate productivity shifts, changes in A, and cost shocks, changes in μ, as equivalent. So what follows does not depend on which of these occurs. We separate the discussion to distinguish between the short- and long-run impact of a decline in S from $S_0 = 0$ to $S_1 < 0$ (lower productivity or higher costs). By the "short-run" we mean for a fixed $m = m_0 = 1/2$. By the "long-run," on the other hand, we mean after *m* has adjusted to its new steady state value $m_1 = 1/2 + (2/3)S_1$.

It is easy to see from (6) and (7) that in the short-run:

$$\frac{\partial D}{\partial S} = -\frac{1}{4} = -\frac{\partial H}{\partial S}. \qquad (9)$$

That is, when S drops, creation falls and destruction rises, leading to a decline in *N* (see (4)). In other words, in a normal economy, negative profit shocks are met with both increased exit by incumbents and reduced entry of new firms.

Over time, the gap between destruction and creation reduces the number of incumbents (recall from (4) and (8) that $\Delta N = H - D$), which lowers the cost of inputs ($p(N)$) and eventually puts an end to the gap between creation and destruction caused by the negative shock.

Across steady states, we have that:

$$\frac{\partial m}{\partial S} = \frac{\partial N}{\partial S} = \frac{2}{3}.$$

The number of production units falls beyond the initial impact as time goes by and the positive gap between destruction and creation closes gradually. Note that since *N* falls less than one for one with S, the long run reduction in the input cost due to reduced competition is not enough to offset the direct effect of a lower S on creation. That is, creation falls in the long run. And since creation and destruction are equal in the long run, the initial surge in destruction is temporary and ultimately destruction also ends up falling below its pre-shock level.[11]

3.5 Zombies

Suppose now that "banks" choose to protect incumbents from the initial surge in destruction brought about by the decline in S. There are a variety of ways that this might be accomplished. We assume that the banks do this by providing just enough resources to the additional units that would have been scrapped so that they can remain in operation.

[11] This long run level effect is undone when creation and destruction are measured as ratios over *N*, as is often done in empirical work. However, the qualitative aspects of the short run results are preserved since empirically the flows are divided by either initial employment or a weighted average of initial and final employment.

With this assumption, a firm that does receive a subsidy is indifferent to exiting and operating, and thus entry and exit decisions remain myopic.

The maximum short run effect would be on impact, when the normal economy would show a spike in destruction (see (5)). Under the zombie-subsidy assumption, we have that:

$$D_{0+}^{z} = D_0 = \frac{1}{4}.$$

The post-shock destruction remains the same as the pre-shock level. The lack of adjustment on the destruction margin means that now creation must do all the adjustment:

$$N_{0+}^{z} = H_{0+}^{z} + m_0 - 1/4 = H_{0+}^{z} + 1/4. \qquad (10)$$

Replacing this expression into (3), we can solve out for *H*:

$$H_{0+}^{z} = \frac{1}{4} + \frac{S}{3},$$

This can be compared to the impact change in creation that occurs in the absence of zombies. Doing so, we see:

$$\frac{\partial H_{0+}^{z}}{\partial S} = \frac{1}{3} > \frac{1}{4} = \frac{\partial H_{0+}}{\partial S}.$$

That is, a decline in S has a much larger negative effect on creation in the presence of zombies. This result is a robust feature of this type of model. In particular, the same qualitative prediction would hold even if we had not suppressed the dynamics and had allowed persistence in the productivity shocks and a gap between entry costs and the productivity advantage of new firms. Intuitively, this is the case because the adverse shock causes the labor market to clear with fewer people employed. If destruction is suppressed, then the labor market clearing can only occur if job creation drops precipitously.

As Caballero and Hammour (1998, 2000) emphasize, both this "sclerosis" — the preservation of production units that would not be saved without the banks' subsidies— and the associated "scrambling" — the retention of firms that are less productive than some of those that do not enter due to the congestion caused by the zombies – are robust implications of models of creative destruction when there are frictions against contracting.

Compared with a normally functioning economy, we have shown the existence of zombies softens a negative shock's impact on destruction and exacerbates its impact on creation. What is the net effect on the number of firms? It is straightforward to show:

$$\frac{\partial N_{0+}^{z}}{\partial S} = \frac{1}{3} < \frac{1}{2} = \frac{\partial N_{0+}}{\partial S}.$$

That is, in response to a negative shock, *N* falls by less if there are zombies, which means that in the presence of zombies the reduced destruction is not fully matched by the additional drop in creation. This is another intuitive and robust result. This occurs because as job creation falls, the marginal entrant's productivity rises. This high productivity allows the marginal entrant to operate despite the higher cost induced by (comparatively) larger *N*.

A final important prediction of the model is the existence of a gap in profitability (net of entry costs) between the marginal entrant and the marginal incumbent when there are zombies.[12] At impact, the destruction does not change, so that all the firms with idiosyncratic productivity shocks above the old threshold (1/2) remain in the industry. On the other hand, new entrants have to clear a higher threshold to compensate for the negative shock in S (which is only partially offset by the lower congestion following the negative shock). As a result, the profitability of the marginal entrant is inefficiently higher than that of the marginal incumbent. The difference is given by:

[12] Note that a wedge like this one also arises when there is a credit constraint on potential entrants but not on incumbents. In our model depressed entry results from the congestion due to zombies, and the gap is due to the subsidy to incumbents. Clearly, however, if the two mechanisms coexist they would reinforce each other, as congestion would reduce the collateral value of potential entrants.

$$\left[\left(\frac{1}{2}+\frac{S_1}{3}\right)-S_1\right]-\frac{1}{2}=-\frac{2}{3}S_1>0.$$

In summary, the model makes two robust predictions. The first is that the presence of zombies distorts the normal creation and destruction patterns to force larger creation adjustments following shocks to costs, productivity or profits. Second, this distortion depresses productivity by preserving inefficient units at the expense of more productive potential entrants. Accordingly, productivity will be lower when there are more zombies and as the zombies become more prevalent they will generate larger and larger distortions for the non-zombies.

By slightly re-interpreting what a "firm" means in our model, we can also see how the congestion effects caused by zombies will affect firms with different levels of profitability. Instead of assuming that a firm has only one project, suppose a firm consists of a set of projects, some of which are in place (incumbents) but the others have not been started (potential entrants). Then, the above model can be re-interpreted as a model in which projects that are hit by productivity shocks every period and firms are deciding which projects to terminate (exits) and which new projects to start (entries).

Suppose further that firms differ in the quality of their projects. In particular, some (high profitability) firms have many projects that are unusually profitable, but some other (low profitability) firms have only a few profitable projects. Low profitability firms will not start many new projects, and the presence of zombies may not influence this very much. Higher profitability firms, however, are more likely to have some new projects that become profitable each period that might be crowded out by the zombies. This effect, however, could be non-monotonic because if a firm has a sufficiently good mix of projects, then its projects might still be worth initiating. We will also test for whether higher quality firms are disproportionately harmed by the zombies, but (because of the potential non-monotonicity) we see this prediction as less robust than the previous two.

4. The effect of zombies on job creation, destruction and productivity

We use the two robust predictions of the model to guide our search for evidence that the zombie problem has affected Japan's economic performance significantly. We begin by looking at aggregate cross-industry differences. In the next section, we study firm-level data to characterize how the behavior of the non-zombie firms has been altered by the presence of zombie competitors.

Because our zombie indices exist from 1981 onwards, we start by calculating the average of the crisp zombie index for each industry from then until 1993 and compare that to the average for the late 1990s (1996-2002). We use the differences in these two averages to correct for possible biases in the level of zombie index and any industry-specific effects. It makes little difference as to how we define the pre-zombie period. In particular, the results we show would be very similar if we took the normal (non-zombie) period to be 1981 to 1990, or 1990 to 1993. Our evidence consists of relating creation, destruction, and productivity data to this change in the zombie index, in order to see if these measures are more distorted in the industries where zombie prevalence has increased the most.

Our most direct evidence on this point is in Figure 5, which plots the rate of job creation and destruction against the change in the zombie index. We use the job flow measures constructed by Genda et al. (2003) as proxies for the concepts of entry and exit in our model. Their measures are based on The Survey of Employment Trends, conducted by the Ministry of Welfare and Labor biannually on a large sample of establishments that employ five or more regular workers. The series used for our analysis include not only the job creation (destruction) at the establishments that were included in the survey in both at the beginning and at the end of the year, but also the estimated job creation (and destruction) by new entrants (and the establishments that exited). To control for the industry specific effects in job creation/destruction, we look at the difference between the average job creation (destruction) rate for the 1996-2000 period and the average for the 1991-1993 period. We are restricted to using the 1991—93 data as a control because figures of Genda et al. start only in 1991 and we stop in 2000 because that is the last year they cover.

The top of Figure 5 shows that the job destruction rate in the late 1990s increased from that in the early 1990s in every industry, as we would expect to see following an unfavorable shock to the economy.[13] More importantly, the graph shows that the surge in destruction was smaller in the industries where more zombies appeared. Thus, as we expected, the presence of zombies slows down job destruction.

The second panel of Figure 5 shows that the presence of zombies depresses job creation. Creation declined more in the industries that experienced sharper zombie growth. In manufacturing, which suffered the least from the zombie problem, job creation hardly changed from the early 1990s to the late 1990s. In sharp contrast, job creation exhibits extensive declines in non-manufacturing sectors, particularly in the construction sector.

Of course not all sectors were equally affected by the Japanese crash in asset prices and the slowdown that followed it. For example, construction, having benefited disproportionately from the boom years, probably also was hit by the largest recessionary shock during the 1990s. A large shock naturally raises job destruction and depresses job creation further. Despite this source of (for us, unobserved) heterogeneity, the general patterns we expected from job flows hold. One way of controlling for the size of the shock is by checking whether in more zombie-affected sectors, the *relative* adjustment through job creation is larger. In this metric, it is quite clear from Figure 5 that job creation has borne a much larger share of the adjustment in construction than in manufacturing.

Our evidence on productivity distortions caused by the interest rate subsidies is given in Figure 6. In the model, zombies are the low productivity units that would exit the market in the absence of help from the banks. Their presence lowers the industry's average productivity both directly by continuing to operate and indirectly by deterring entry of more productive firms. The productivity data here are from Miyagawa, Ito and Harada (2004) who study productivity growth in 22 industries. Figure 6, which plots the average growth of the total factor productivity (TFP) from 1990 to 2000 against the

[13] Our simple model assumes that the job destruction rate stays the same even after a negative shock in a zombie industry. It is straightforward to relax this by assuming, for example, that 90% of zombies are rescued by banks. None of the major results would change. Job destruction would rise following a negative shock but not as much as it would under the normal environment.

change in the crisp zombie index, shows that the data are consistent with the model's implication: the regression line in the figure confirms the visual impression that industries where zombies became more important were the ones where TFP growth was worst.[14]

5. Firm-level zombie distortions

We read the evidence in Figures 5 and 6 as showing that zombies are distorting industry patterns of job creation and destruction, as well as productivity in the ways suggested by the model. To test directly the model's predictions, we next look at firm-level data to see if the rising presence of zombies in the late 1990s had discernible effects on the healthy firms (which would suffer from the congestion created by the zombies).

The data we analyze are from the Nikkei Needs Financial dataset and are derived from income statements and balance sheets for firms listed on the first and second sections of the Tokyo Stock Exchange. The sample runs from 1981 to 2002, and it contains between 1,844 and 2,506 firms depending on the year. We concentrate on three variables: employment growth (measured by the number of full-time employees), the investment rate (defined as the ratio of investment in depreciable assets to beginning of year depreciable assets measured at book value), and a crude productivity proxy (computed as the log of sales minus 1/3 the log of capital minus 2/3 the log of employment). In all the regressions reported below we dropped observations in the top and bottom 2.5% of the distribution of the dependent variable.

The simplest regression that we study is:

$$\text{Activity}_{ijt} = \delta' D_{jt} + \beta \text{nonz}_{ijt} + \chi Z_{jt} + \varphi \text{nonz}_{ijt} * Z_{jt} + \varepsilon_{ijt} \qquad (11)$$

where activity can be either the investment rate, the percentage change in employment, or our productivity proxy, D_{jt} includes a set of annual indicator variables and a set of industry dummy variables, nonz_{ijt} is the probability that the firm is non-zombie, and Z_{jt} is the percentage of industry assets residing in zombie firms.

---

[14] Of course this correlation could arise because industries that had the worst shocks wound up with the most zombies. We can disentangle these explanations by using firm-level data (see below).

Because of the reduced form nature of both the regression equation and the modeling of the subsidies to the zombies, we do not attempt to interpret most of the coefficients in these regressions. For instance, we include the year dummies to allow for unspecified aggregate shocks. Likewise, we can imagine that the zombies' subsidies are so large that they wind up investing more (or adding more workers) than the healthy firms; so we do not propose to test the theory by looking at the estimates for β, the coefficient for the non-zombies. The one exception to this general principle is that for the productivity specification the model clearly predicts that non-zombies will have higher average productivity than zombies.

We instead focus on what we see as the novel prediction of the theory: that the rising zombie congestion should harm the non-zombies. This prediction suggests that φ would be negative in the investment and employment regressions, and be positive in the productivity specification. Note that for the investment (employment) specification one might normally suspect that as the percentage of sick firms in the industry rises, the healthy firms would have more (relative to the sick ones) to gain from investing (expanding employment). Thus, under normal circumstance there would be good reasons to expect φ to be positive rather than negative.

The main reason, other than ours, for finding negative φ is if the zombie percentage in the industry is somehow standing in for the overall (un)attractiveness of operating in the industry. To this potential objection to our results we note two things. First, our definition of zombies, by virtue of only using interest rate payments, does not guarantee that growth opportunities are necessarily bad just because the zombie percentage is high. Second, in order to be consistent with our findings, the reaction to industry conditions must be different for zombies and non-zombies. In particular, non-zombies must be more affected by an industry downturn than zombies for φ to come out negative. Nonetheless, we seek to find other controls for business opportunities for the healthy firms to minimize this potential omitted variable bias. Our main control to address this problem is to add current sales growth of each firm to the regression specification. Thus, our alternative regression is:

$$\text{Activity}_{ijt} = \delta' D_{jt} + \beta \text{nonz}_{ijt} + \chi Z_{jt} + \varphi \text{nonz}_{ijt} {*} Z_{jt} + \theta s_{ijt} + \psi \text{nonz}_{ijt} {*} s_{ijt} + \zeta s_{ijt} {*} Z_{jt} + \pi \text{nonz}_{ijt} {*} Z_{jt} {*} s_{ijt} + v_{it} \qquad (12)$$

where s_{ijt} is the growth rate of sales and the other variables are defined as in equation (11).

The coefficient π in (12) reveals an additional potential effect for the zombies. If π is different from zero, then it implies that faster growing healthy firms and slower growing healthy firms are differentially affected by the presence of the zombies. As mentioned earlier, a natural interpretation of the model suggests that the zombie distortions should be larger for the healthiest firms. This would be the case if $\pi < 0$.

The second through fourth columns of Table 2 shows our estimates for equations (11) for the crisp zombie index. We draw three main conclusions from this simple specification. First, as predicted by the theory, increases in percentages of zombie firms operating in an industry significantly reduces both investment and employment growth for the healthy firms in the industry.[15] Our second finding, shown in column 4, is that the non-zombies have significantly higher productivity than the zombies. Finally, the same column shows that the productivity gap between zombies and non-zombies rises as the percentages of zombies in an industry rises. These findings are consistent with the main predictions of our model.

As mentioned above, a competing explanation for the sign of the estimated φ in equation (11) is that the industry zombie percentage is an indirect measure of the growth opportunities in the industry, even for the healthy firms. We address this concern by including controls that directly capture growth opportunities. Columns 5 and 6 report estimates of equation (12), which include contemporaneous firm-specific sales growth as the potential growth proxy; for the investment specification, this type of accelerator specification generally performs quite well in a-theoretic horse-races among competing specifications (see Bernanke, Bohn and Reiss (1988)).

[15] We ran a similar regression using investment rates for US firms covered in the Compustat database between 1995 and 2004. In this regression φ was insignificantly different from zero. The limited information on debt structure in Compustat no doubt introduces noise in zombie assignments and we did explore many alternatives to deal with this. But this result suggests to us that there is not a mechanical reason to find that φ is significantly negative in this type of regression.

In both columns the estimated coefficient on sales growth is highly significant, and in each equation the $\overline{R}^2$ is nearly twice as high as that in the simpler specifications in columns 2 and 3. In the specifications with sales growth, the estimated magnitude of the φ's drops compared to the simpler specifications, but they remain negative and significant. This indicates that while some of the interaction term's significance may have been due to omitting proxies for growth opportunities, it is not the sole reason.

More substantively, in both of these specifications the estimated values for π are significantly negative. This triple interaction suggests that the fastest growing non-zombie (healthy) firms are the most impaired by the widespread presence of zombie firms in their industry.

In Appendix 2 we report a long list of robustness exercises, including fuzzy versions of equations (11) and (12), regressions omitting marginal zombies, as well as using different measures of minimum required interest rates in the construction of zombie indicators. While the level of significance and some of the point estimates vary across these multiple scenarios, the general flavor of the results does not. More specifically, the estimates for φ tend to be negative and significant for the investment and employment regressions and positive and significant for the productivity regressions. The estimates of π are more sensitive to the exact specification, and vary more for the employment regressions than for the investment specifications.

In the remainder of our discussion we attempt to quantify the impact of zombie firms on investment and employment growth of non-zombies. We focus on the five non-manufacturing industries, where our asset weighted measures of zombies were particularly high in the late 1990s. For a typical non-zombie firm in each of these industries, we estimate how much more the non-zombie would have invested or increased employment if there had not been so many zombies in the industry. We consider two alternative low zombies scenarios. In "Case 1," we assume that the zombie index stayed at its average value from 1981 through 1992 for each industry and calculate how much more a typical non-zombie firm would have invested (or employed) over the next ten years.[16] In "Case 2," we assume that the zombie index for the industry was the same as

[16] More specifically, the investment (or employment) is estimated to have been higher than the actual level by $(\hat{\chi} + \hat{\varphi})(actual \quad zombie \quad index - alternative \quad zombie \quad index)$.

that for manufacturing for each year from 1993 to 2002. We calculate the cumulative investment under these two scenarios and compare it to the typical amount of annual investment (defined as the average of the median rates) during this period. For employment, we compare the cumulative decline attributable to the zombies with the typical annual change over the period (again defined as the average of the median rates). In all of these calculations we take the regression estimates based on the crisp zombie indices in Table 2, and ignore any feedback from industry equilibrium considerations.

Table 3 shows both investment and employment growth in non-zombie firms would have been higher in all these industries had there been less zombies. In some industries, the difference is quite large. For example, for the typical non-zombie firm in the wholesale industry the cumulative investment loss was about 12.1% of capital, which was slightly more than one year worth of investment during this period. The employment growth of a typical non-zombie real estate developer would have been about higher by 3.0 percentage points at the end of the period if the zombie percentage had not risen (which can be compared to the average hiring in the industry of 0.62% per year). Overall, these effects are substantial.

In our main specifications we find the effect of zombie infestation on non-zombies depends on the level of sales growth of the non-zombie (negative coefficient estimates on the three way interaction). While these triple interaction results are less robust than the double interactions, it is still interesting to document the magnitude of the differential impacts suggested by our estimates.

Figure 7 uses estimates from Table 2 for equation (12) to infer the differential effect of varying degrees of zombie infestation for non-zombies with different levels of sales growth; formally, this amounts to studying $\frac{\partial^2 Activity}{\partial nonz \partial Z} = \phi + \pi s$. The left panel shows the zombie distortion on investment is significantly worse for fast growing firms; the dotted lines in the graph show the 95 percent (asymptotic) confidence intervals. Not only are these marginal effects significant, the overall quantitative impact is large. For instance, for a firm with ten percent sales growth, if the industry zombie percentage were to increase from 0.1 to 0.2, investment would fall by 1.3 percentage points per year; if the firm instead had 15 percent annual sales growth, the investment drop would be 1.55

percentage points per year. Given the median investment rate of 14.7% per year (1993-2002) we view these effects as large.

The right panel in the figure presents an analogous calculation for employment. The marginal effects again are significant (for all cases where sales growth is above two percent per year). For a firm with sales growth of ten percent per year, an increase in the zombie percentage from 0.1 to 0.2 would depress annual employment growth by 0.25 percentage points per year. Since employment growth for this sample of firms was approximately zero, the implied cumulative effect of the high level of zombies during the late 1990s is big.

Given the depressed condition of the economy between 1993 and 2002 it is not clear which benchmark to use in gauging the size of the effects. Normally, we would have expected to find some firms with sales growth of 10 to 15 percent per year, but these firms are quite rare in Japan over this period. Nonetheless, it appears that there were substantial distortions for the healthiest firms.

6. Final Remarks

Let us now take stock and discuss the implications of our empirical findings. First, the mechanism we have highlighted compounds the problems caused by a traditional credit crunch. Recall that the reduced form profit shock that we analyze in the model subsumes a simple credit crunch. Thus, if a pure contraction in credit availability was all that was going on, the economy would be expected to behave like the normal benchmark case we analyze. It follows that the evidence we presented to support the zombie model, also shows that a pure credit crunch explanation (a la Kitasaka and Ogawa (2000)) for the recent experience, while highly relevant, is insufficient.[17]

One key characteristic of our mechanism is that zombies create on-going distortions that lower job creation and industry productivity. A straightforward extension

[17] There are also other implications that we have not tested that could be further used to distinguish these two models. For instance, the zombie model explains why the firms that do enter or expand need not have high values of Tobin's Q – essentially because the zombie congestion costs lower their profitability. In contrast, a standard credit crunch model would predict that these firms should be earning rents by virtue of being able to operate against reduced competition. See Caballero and Hammour (2005) for a discussion of the channels through which financial factors may depress restructuring during recessions.

of the model would make long-run productivity growth endogenous. In this case the present value of the costs due to the suppression of restructuring generated by continuing forbearance with the zombies would greatly exceed calculation based only on the direct costs of subsidies.

Japanese regulators may have failed to recognize the large costs of allowing zombies to continue operating during the episode. For example, the capital injections given to Japanese banks in the late 1990s did not recapitalize the banks sufficiently so that they no longer had an incentive to evergreen. The forgone benefits that would have accrued had Japan returned at that point to having a normally functioning economy could have been large enough to justify a very generous transition policy package to the displaced workers that would have been released if the zombies were shuttered.[18]

Finally our description of the Japanese experience is similar to the diagnosis that has been used to describe the early phases of the transition of many former socialist economies to becoming market-oriented. In these economies the depressing effects on the private sector of the continued operation of state-owned enterprises (typically funded by state owned banks) is often noted; discussions of the current situation in China would be the latest of these examples. Also, note that the key to our mechanism is lack of restructuring, which may be also caused by legal bankruptcy procedures that protect debtors rather than by banks' behavior. For example, in the U.S. airline industry it is routinely asserted that the industry has been plagued because unprofitable carriers go bankrupt, yet they fail to exit the industry (see Wessel and Carey (2005)). These cases suggest that the mechanism that we have sketched is not unique to Japan.[19]

[18] The same reasoning applies to the question of whether the lack of liquidations in the U.S. airline industry raised or lowered the taxpayers' costs of rationalizing the industry.
[19] See Caballero (2006) for a discussion of different models and manifestations of sclerosis in macroeconomics.

References

Ahearne, Alan G., and Naoki Shinada, 2004, "Zombie Firms and Economic Stagnation in Japan," paper presented at the conference "Macro/Financial Issues and International Economic Relations: Policy Options for Japan and the United States" at University of Michigan, Ann Arbor, MI, October 22-23, 2004.

Bernanke, Ben S., Bohn, Henning, and Peter C. Reiss, 1988, "Alternative Non-Nested Specification Tests of Time Series Investment Models," *Journal of Econometrics*, 37(2), March, pp. 293-326.

Caballero, Ricardo J., 2006, *The Macroeconomics of Specificity and Restructuring,* Yrjo Jahnsson Lecture, forthcoming.

Caballero, Ricardo J. and Mohamad L. Hammour, 1998, "The Macroeconomics of Specificity," *Journal of Political Economy,* 106(4), pp. 724-767.

Caballero, Ricardo J. and Mohamad L. Hammour, 2000, "Creative Destruction and Development: Institutions, Crises, and Restructuring," *Annual World Bank Conference on Development Economics 2000,* pp. 213-241

Caballero, Ricardo J. and Mohamad L. Hammour, 2005, ``The Cost of Recessions Revisited: A Reverse Liquidationist View,'' *Review of Economic Studies*, 72, 313-341.

Fukao, Mitsuhiro, 2000, *Kin'yū Fukyō no Jisshō Bunseki (Empirical Analyses of Financial Recession),* Tokyo: Nihon Keizai Shimbun-sha, (in Japanese).

Genda, Yuji, Hiroshi Teruyama, Soichi Ohta, Ryo Kanbayashi, Mamiko Ishihara, Yuji Senuma, Kazuhiro Sasaki, Kentaro Abe, Takayuki Kusajima, and Taku Morifuji, 2003, "Koyō Sōshutsu to Shitugyō ni kansuru Jisshō Kenkyū" (Job Creation and

Unemployment: Empirical Analysis) *Keizai Kenkyu*, No.168, ESRI (Economic and Social Research Institute), Cabinet Office (Japanese Government), (in Japanese).

Guner, Nezih, Ventura, Gustavo, and Xu Yi, 2004, "Macroeconomic Implications of Restrictions on Size", working paper, Pennsylvania State University.

Hamao, Yasushi, Mei, Jianping, and Yexiao Xu, forthcoming, "Unique Symptoms of Japanese Stagnation: An Equity Market Perspective," *Journal of Money, Credit and Banking*.

Hoshi, Takeo, 2006, "Economics of the Living Dead", *Japanese Economic Review*, 57(1), pp. 30-49.

Hoshi, Takeo, 2000, "Naze Nihon wa Ryūdōsei no Wana kara Nogarerareainoka? (Why is the Japanese Economy Unable to Get Out of a Liquidity Trap?)" in Mitsuhiro Fukao and Hiroshi Yoshikawa eds. *Zero Kinri to Nihon Keizai (Zero Interest Rate and the Japanese Economy)*, Tokyo Nihon Keizai Shimbunsha, pp. 233-266, (in Japanese).

Hoshi, Takeo and Anil K Kashyap, 2001, *Corporate Financing and Governance in Japan*, MIT Press: Cambridge.

Hoshi, Takeo and Anil K Kashyap, 2004, "Japan's Financial Crisis and Economic Stagnation" *Journal of Economic Perspectives*, 18(1), Winter 2004, pp. 3-26.

Hoshi, Takeo and Anil K Kashyap, 2005, "Solutions to Japan's Banking Problems: What Might Work and What Definitely Will Fail," in Takatoshi Ito, Hugh Patrick, and David E. Weinstein (Eds.) *Reviving Japan's Economy: Problems and Prescriptions.* Cambridge, MA: MIT Press, pp.147-195.

Hosono, Kaoru, and Masaya Sakuragawa, 2003, "Soft Budget Problems in the Japanese Credit Market," Nagoya City University Discussion Papers in Economics No.345.

Ito, Takatoshi and Yuri N. Sasaki, 2002, "Impacts of the Basle Capital Standard on Japanese Banks' Behavior," *Journal of the Japanese and International Economies*, 16, pp. 372-397.

Jerram, Richard, 2004, "Banks still not reflating", ING Financial Markets, June 2, 2004.

Kim, Se-Jik, 2004, "Macro Effects of Corporate Restructuring in Japan", International Monetary Fund, working paper.

Kitasaka, Shin-Ichi and Kazuo Ogawa, 2000, "Bank Lending in Japan: Its Determinants and Macroeconomic Implications," in *Crisis and Change in the Japanese Financial System*, eds. Hoshi,T. and H.Patrick, Kluwer Academic Publishers, 2000, pp.159-199.

Miyagawa, Tsutomu, Ito, Yukiko and Nobuyuki Harada, 2004, "The IT revolution and productivity growth in Japan", *Journal of the Japanese and International Economies*, 18(3), pp. 362-389.

Nakakuki, Masayuki, Akira Otani, and Shigenori Shiratsuka, 2004, "Distortions in Factor Markets and Structural Adjustments in the Economy", Bank of Japan Discussion Paper No. 2004-E-4.

Nguyen, Hung T., and Elbert A. Walker, 2006, *A First Course in Fuzzy Logic*, 3rd edition. Boca Raton, FL: Chapman & Hall.

Nishimura, Kiyohiko, Takanobu Nakajima, and Kozo Kiyota, 2003, "Ushinawareta 1990-nen dai, Nihon Sangyō ni Nani ga Okottanoka?" (What Happened to the Japanese Industry in the Lost 1990s?), RIETI Discussion Paper 03-J-002.

Peek, Joe and Eric S. Rosengren, 2005, "Unnatural Selection: Perverse Incentives and the Misallocation of Credit in Japan", *American Economic Review*, 95(4), pp. 1144-1166.

Restuccia, Diego and Richard Rogerson, 2003, "Policy Distortions and Aggregate Productivity with Heterogeneous Plants", working paper, Arizona State University.

Sakai, Koji, Iichiro Uesugi, and Tsutomu Watanabe, 2005, "Firm Age and the Evolution of Borrowing Costs: Evidence from Japanese Small Firms," RIETI Discussion Paper 05-E-026.

Sakuragawa, Masaya, 2002, *Kin'yū Kiki no Keizai Bunseki (Economic Analysis of Financial Crisis),* Tokyo: Univeristy of Tokyo Press, (in Japanese).

Sasaki, Yuri, 2004, "Japanese Banks' Behavior under the Basle Capital Standards, mimeo, Meiji Gakuin University.

Schaede, Ulrike, 2005, "The 'Middle Risk Gap' and Financial System Reform: Small Firm Financing in Japan," *Monetary and Economic Studies*, 23 (1), pp.149-176.

Sekine, Toshitaka, Kobayashi, Keiichiro, and Yumi Saita, 2003, "Forbearance Lending: The Case of Japanese Firms", *Monetary and Economic Studies*, 21(2), pp. 69-91.

Smith, David C., 2003, "Loans to Japanese Borrowers", *Journal of the Japanese and International Economies*, 17(3), pp. 283-304.

Tett, Gillian, 2003, *Saving the Sun : A Wall Street Gamble to Rescue Japan from Its Trillion-Dollar Meltdown*, Harper-Collins Business.

Tett, Gillian and David Ibison. 2001. "Tokyo 'May Have to Support Banks'." *Financial Times*, September 14.

Wessel, David and Sarah Carey, 2005, "For U.S. Airlines, a Shakeout Runs Into Heavy Turbulence", *Wall Street Journal*, September 19.

White, Halbert, 1980, “A Heteroscedasticity-Consistent Covariace Matrix Estimator and a Direct Test for Heteroscedasticity,” *Econometrica*, 48, 817-838.

Appendix 1

The variable R^* plays a critical role in our analysis. In this appendix we provide some additional details on the construction of this variable and the other data used in the analysis.

In constructing R^* our goal is to produce a plausible lower bound for what firms might pay to borrow. For the portion of the interest payments coming from short term bank loans, which accounts for about 40% to 45% of total lending in our sample, we believe that this is straightforward because almost no loans are made at rates below the prime rate (once we take into account all the origination and other fees). Thus, we view the use of the short term prime rate as relatively uncontroversial.[20]

Ideally, we would find an equally conservative assumption for handling long-term loans. It is quite likely that interest payment on a new long-term loan would be above the prime rate at the time the loan is originated. Unfortunately, the available data on long-term bank debt gives just the stock outstanding without information on the exact maturity of the loans. So we assume that each firm's long term loans have an average maturity of 2.5 years and with one-fifth having been originated in each year for five years. Five years corresponds to the average maturity of bank loans in the dataset of Smith (2003). This assumption implies that the right interest rate is an equally weighted average of the last five years of the long-term prime rates. Thus, we calculate the minimum required interest payment on the long-term loans by multiplying the outstanding long-term loans of all maturities with the five year average of the long-term prime rates.

Turning to the non-bank financing, we know that during the 1990s, roughly 40% of interest paying debt was bonds and about 3% was commercial paper. Our measure of the required payment ignores the interest payments for commercial paper. Given the limited importance of commercial paper financing and the low interest rates on the commercial paper for the 1990s, this is not likely to cause any serious problems for our analysis.

[20] As alternative we instead computed a required rate that imposed a mark up over the London Interbank Borrowing (LIBOR) rate based on the average spreads reported in Smith (2003). This approach produced similar results regarding the numbers of firms with negative interest rate gaps.

For the remaining debt we assume that it was financed as advantageously as possible. Specifically, we assume that bond financing is done with CBs (which by their nature have lower yields) and that firms were always able to time the issues so that the rate is the lowest within the last five years. Implicitly, this presumes that the firms have perfect foresight and refinance their bonds every time there is a local trough in interest rates. This assumption is almost surely going to understate the required payments on corporate debt. For instance, from 1996 onwards this imputation procedure will assume that all bond financing is done at a zero interest rate. By assuming very low required interest rates on bonds, the approach reduces the risk of our misclassifying credit worthy companies that enjoy extreme low bond rates in the public market as zombies. On the other hand, the approach would increase the risk of failing to identify the zombies that pay interests on the bonds they issued in the past. Thus, we can be confident that any firms that we label as zombies must be getting very favorable interest rates from their banks. Put differently, by assuming access to such low bond financing rates our classification scheme will pick out only the most egregious zombies that receive massive help from their banks.

Besides this baseline procedure we also explored several approaches. One alternative centered on estimating the maturity structure of each firm each year. Here we just describe the calculation for long-term bank borrowing. We estimate the maturity structure of bonds in the same way.

We observe the total long-term bank borrowing for firm i at the end of accounting year t (BL_{it}) and the long-term bank borrowing that comes due within 1 year ($BL1_{it}$). Let NBL_{it} be the amount of new long-term bank loans that the firm i takes in during year t. We use the following equation to estimate NBL_{it}:

$$NBL_{it} = \max\left\{BL_{it} - BL_{it-1} + BL1_{it-1}, 0\right\}$$

Let $BP(n)_{it}$ denote the amount of long-term bank loans to firm *i* that was given in year *t*-*n* and still outstanding at the end of *t*. We assume the maximum maturity of long-term bank loans to be 10 years. If *NBL* is available for all years in the past 10 years, we can estimate *BP(n)* recursively as follows.

$$BP(0)_{it-1} = \min\left\{NBL_{it-1}, \max\left\{BL_{it-1}, 0\right\}\right\}$$

$$BP(n)_{it-1} = \min\left\{NBL_{it-n-1}, \max\left\{BL_{it-1} - \sum_{k=0}^{n-1} BP(k)_{it-1}, 0\right\}\right\} \quad (n = 1, 2, \cdots, 8)$$

$$BP(9)_{it-1} = \max\left\{BL_{it-1} - \sum_{k=0}^{8} BP(k)_{it-1}, 0\right\}$$

If NBL_{it-n-1} is not available for $n \geq n^*$, we stop the iteration at $n = n^*$ and assume that the remaining borrowings (if any) are uniformly distributed across different maturities. Formally, this implies:

$$BP(0)_{it-1} = \min\left\{NBL_{it-1}, \max\left\{BL_{it-1}, 0\right\}\right\}$$

$$BP(n)_{it-1} = \min\left\{NBL_{it-n-1}, \max\left\{BL_{it-1} - \sum_{k=0}^{n-1} BP(k)_{it-1}, 0\right\}\right\} \quad (n < n^*)$$

$$BP(n)_{it-1} = \max\left\{\frac{BL_{it-1} - \sum_{k=0}^{n^*-1} BP(k)_{it-1}}{10 - n^*}, 0\right\} \quad (n \geq n^*)$$

The associated regression results are shown in Table A-3 (that we discuss in Appendix 2).

For bonds, we also adopted an extremely conservative approach that assumes the minimum required interest rate for bonds was zero for the entire sample period. This approach guarantees that any firms with a negative interest rate gap must be receiving unusually low interest rates on their bank borrowing. The regressions associated with this classification scheme are shown in Table A-4 (and are almost identical to those shown in Table 2).

The data for prime bank loan rates are taken from the Bank of Japan web site (http://www.boj.or.jp/en/stat/stat_f.htm). The subscribers' yields for convertible bonds are collected from various issues of *Kin'yu Nenpo* (Annual Report on Finance) published by the Ministry of Finance. The remaining data we use for the regression analyses are taken from the Nikkei Needs Corporate Financial Database. The data are annual, so for instance when we refer to 1993 data they are from a firm's balance sheet and income statement for the accounting year that ended between January and December of 1993.

Appendix 2

We checked the robustness of the significance of the estimated φ's and π's to several alternative measures of the required minimum interest rate r^* and zombie indices. Table A-1 repeats the regressions from Table 2, using the fuzzy zombie indices with $(d_1, d_2) = (0, 50bp)$ and $(d_1, d_2) = (-25bp, 75bp)$. We draw three conclusions from this table. First, the estimates of φ and π are smaller than those in Table 2. However, part of the difference can be explained by the fact that the industry zombie percentages are larger when we use the fuzzy zombie measures than when we use the crisp measures. Second, and probably related, for the estimates of (11), the statistical significance of the estimates of φ is similar to those in Table 2; in other words, the declines in the size of the coefficients are accompanied by smaller standard errors, so that the t-statistics are similar.

When we add sales growth (and the associated interaction terms) to the equations the significance of both φ and π falls. Their estimated signs remain negative in all cases, but most of the coefficients are no longer significant.[21] So this specification is less robust to this alternative measure of zombies.

We also estimated the regressions dropping the observations with x_{it} between d_1 and d_2 entirely. Table A-2 shows the results. The estimates of φ in the investment and employment growth equations are again negative and statistically significant in almost all the cases. Indeed, the size of the coefficient is often higher when we drop the observations with x_{it} close to zero. For the productivity proxy, however, the estimated gap between the zombies and non-zombies (β in equation 11) rises substantially, while the estimated value of φ falls dramatically and becomes insignificant. The result for the three-way interaction term (non-zombie dummy x sales growth x industry zombie percentage) is not robust to this change in specification, either, suggesting the result critically depends on the inclusion of the observations with x_{it} close to zero.

We considered several other alternatives that are not reported since the results are so similar to those shown in Tables 2, A-1, and A-2. In particular, the regressions in Tables 2, A-1 and A-2 consider only the post 1993 period, when the zombie percentages

[21] We could in principle use a mixture of crisp and fuzzy assignments to separate the individual firm classifications and the industry zombie percentages. We performed a few experiments of this type for employment growth and did not find any systematic ways in which the results were affected.

began to rise noticeably. When we re-estimated the Table 2 regressions to include the 1980s in the sample, the estimates for investment and employment growth remain unchanged, while those for productivity change. The estimated gap between the zombies and non-zombies rises substantially, while the estimated value of φ falls sharply and becomes insignificant.

We also tried different definitions of non-zombies. Specifically, we counted a firm as a non-zombie only if it is not classified as a zombie in two or three consecutive years. In both cases the estimates of φ continued to be significantly negative for the investment and employment regressions, and significantly positive for the productivity regression, but for one exception where the estimate of φ, while still negative, was not significant even at 20% level. The estimates of π were never statistically significant with this alternative definition of non-zombies, although the point estimates remained negative.

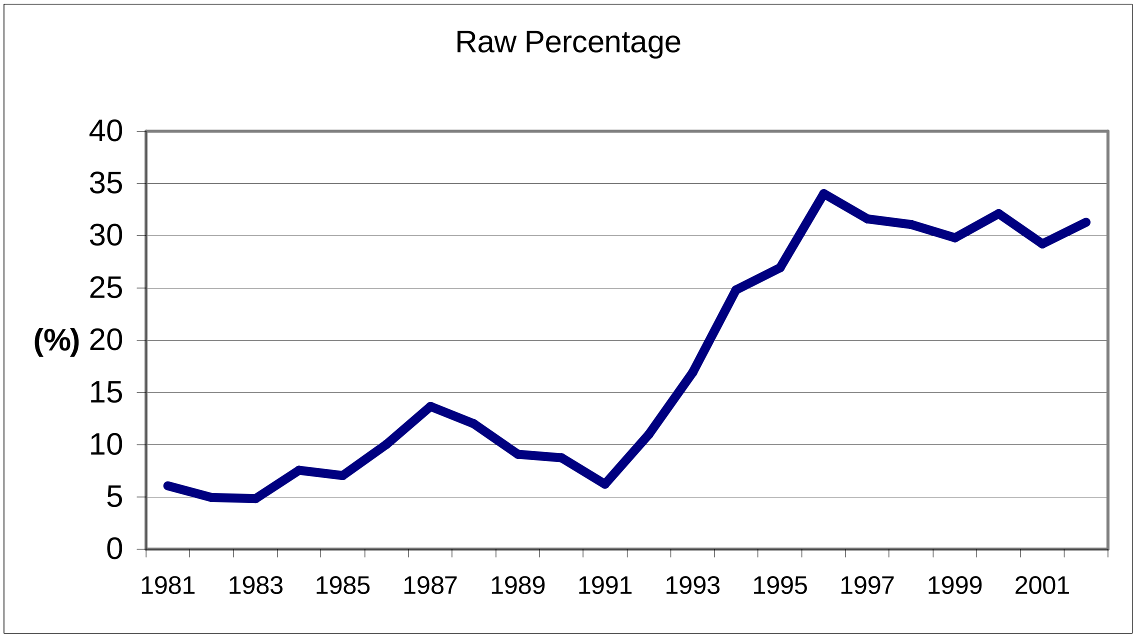
This alternative way of defining non-zombies can be applied to the fuzzy zombie indices as well. We did this by recoding the zombiness of each firm in each year to be the maximum of the z calculated using the equation (1) over the last two (or three) years. Thus, to be classified as a non-zombie for sure, a company has to have $z = 0$ for 2 (or 3) consecutive years. The regression results did not differ much from those in Table A-2. The estimates of φ are statistically significant with expected signs in the regressions without the sales growth. With sales growth, the estimates of π are not significant, and the estimates of φ often lose significance, although the point estimates remain negative.

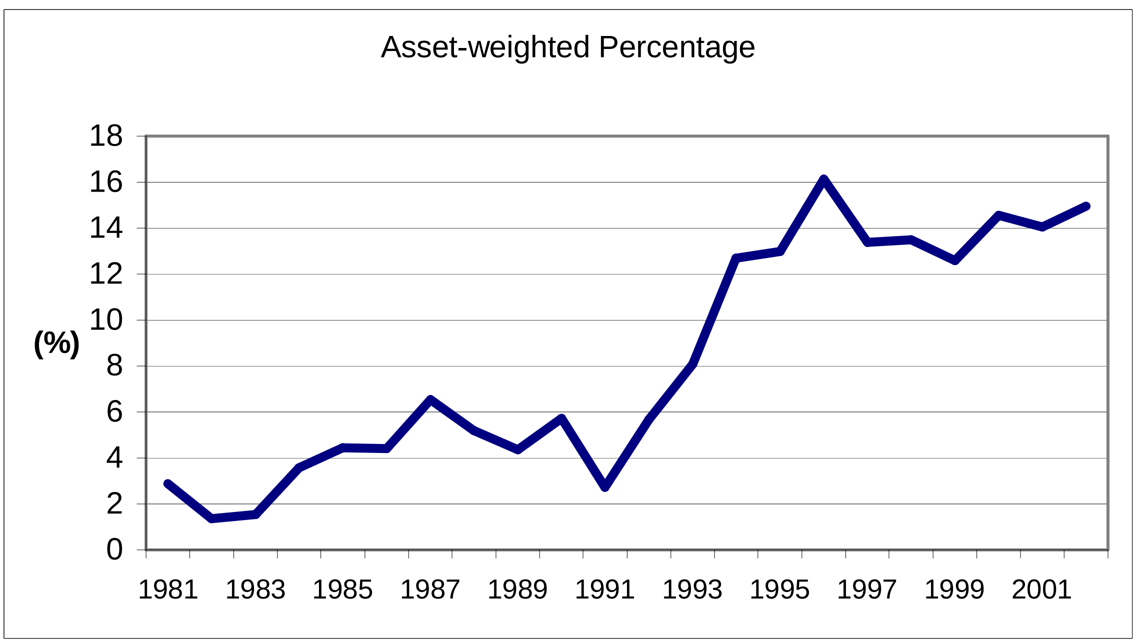
Table A-3 shows the results using more detailed estimation of the maturity structure for long-term borrowings and bonds discussed in Appendix 1. The coefficient estimates on the simple interaction term (non-zombie dummy times industry zombie percentage) are similar to those in Table 2 in all the specifications. The standard errors are, however, sometimes larger, so that the estimates are statistically significant only at 10%. Under this alternative assumption about the maturity structure, the results for the three way interaction term (non-zombie dummy times sales growth times industry zombie percentage) disappear. The coefficient estimates on the three way interaction in the last two columns are not significantly different from zero.

Finally, Table A-4 shows the regressions under alternative assumption that the minimum required interest rate on bonds is zero. The results are essentially the same as those in Table 2 except for the last column. In the employment change equation with sales growth variables, the estimates of the interaction terms cease to be statistically significant, although the point estimates fall only by small amounts.

All in all, the results of these robustness exercises are consistent with those in the main text, although it is apparent that the precisions of some of our estimates suffer as we dilute the measures of zombism and increase their robustness to different measurement and classification errors.

Figure 1: Prevalence of Firms Receiving Subsidized Loans in Japan





Note: Percentages calculated as described in the text, with d_1=d_2=0 in equation 1.

Figure 2: Membership Function for a Fuzzy Zombie Set

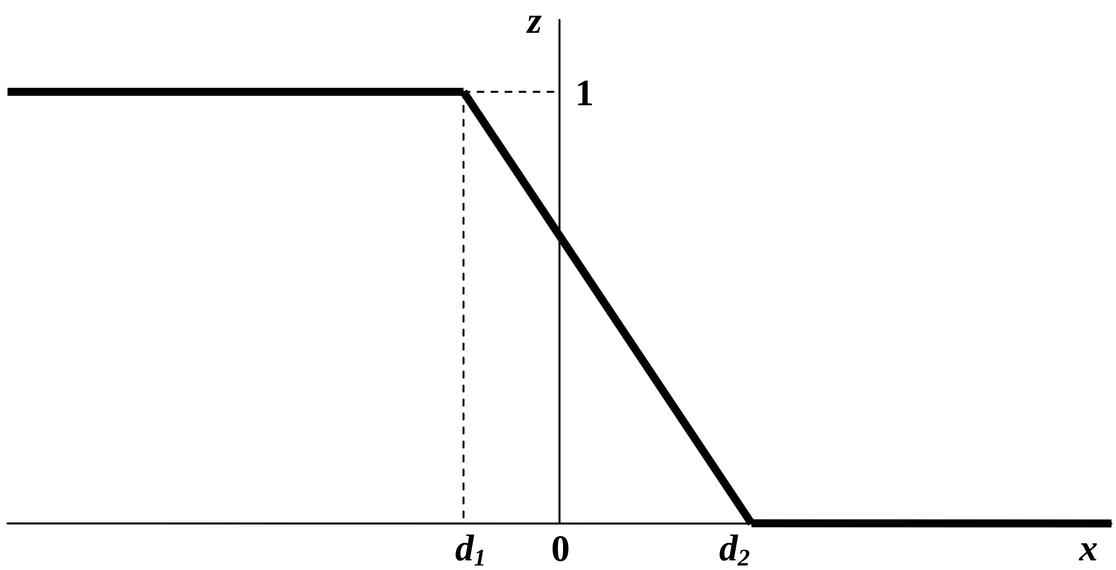
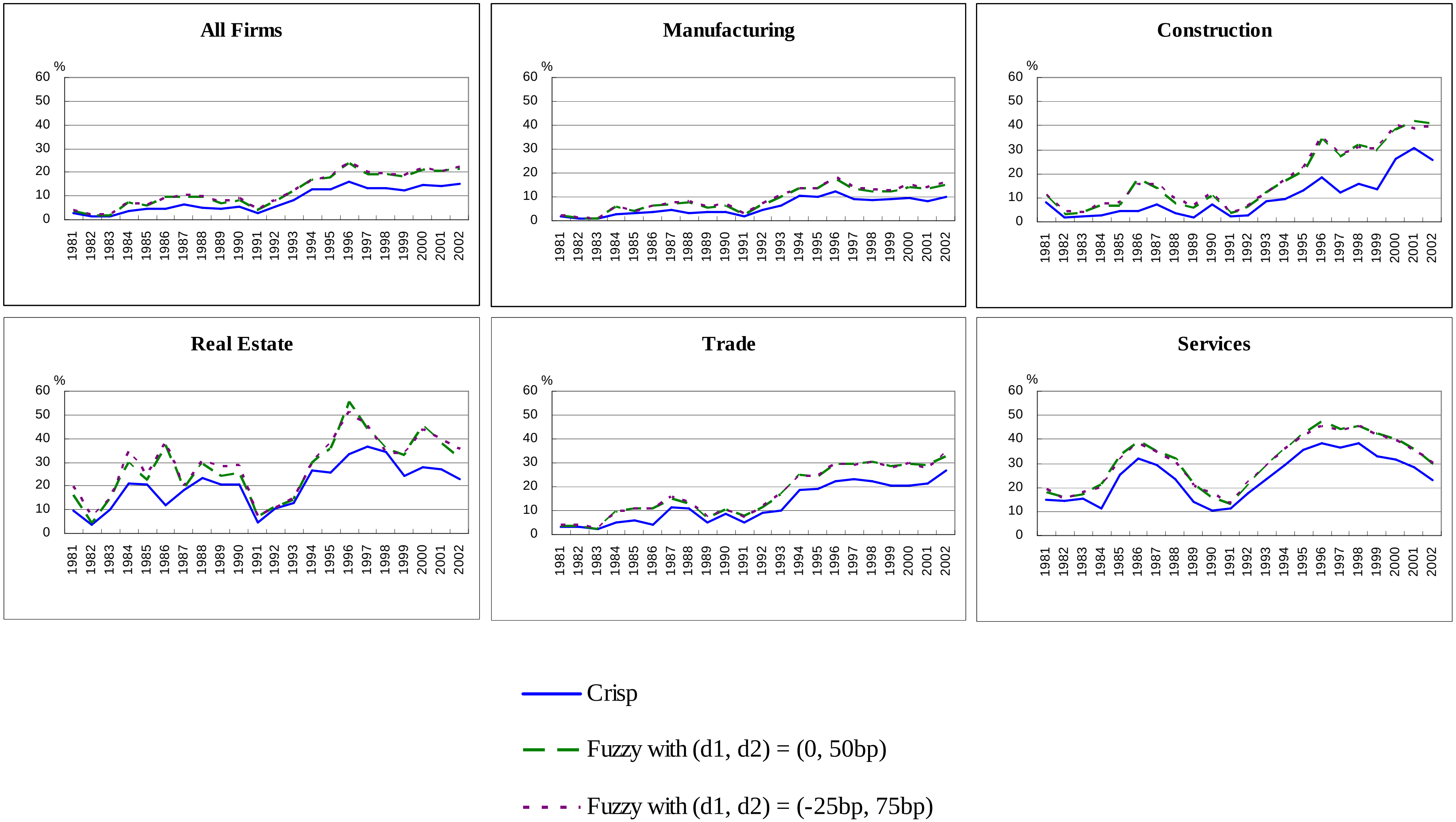
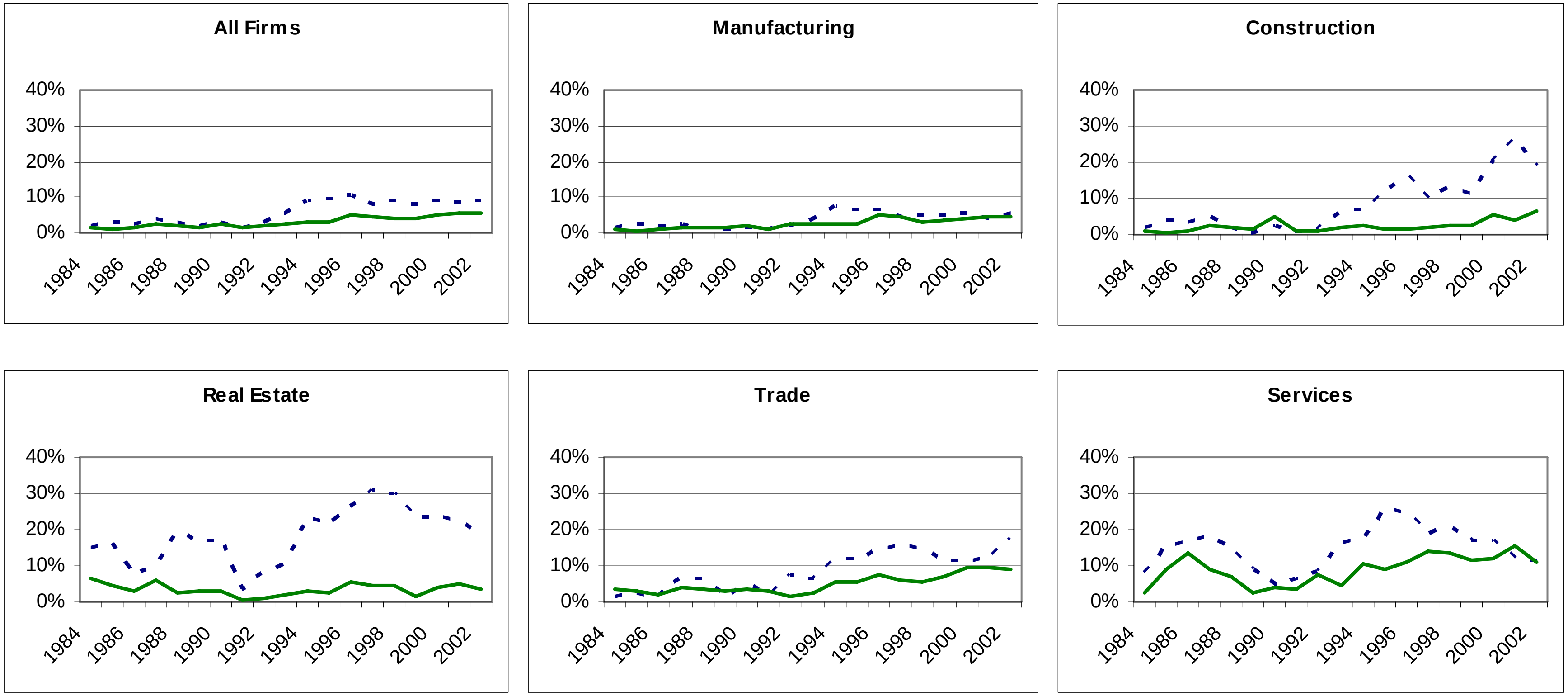


Figure 3: Cross-Industry Incidence of Asset Weighted Zombie Percentage for Crisp and Fuzzy Zombie Definitions



Note: Fuzzie zombie definitions computed according to equation 1, see text for details.

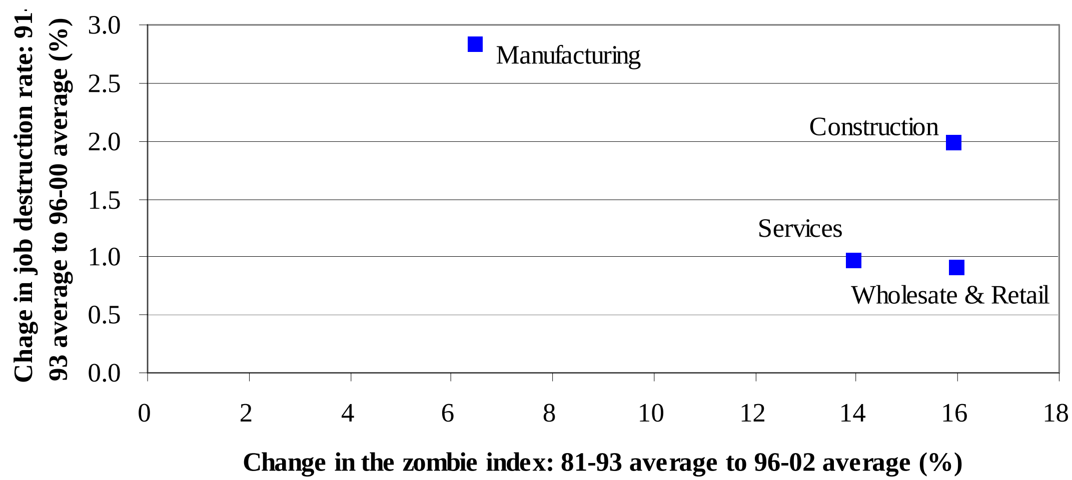
Figure 4: Asset Weighted Zombie Percentages by Profitability



Note: Solid lines show zombie percentage for firms whose profits are above the median for the industry, dashed show below median.

Figure 5

Zombies and Job Destruction



Zombies and Job Creation

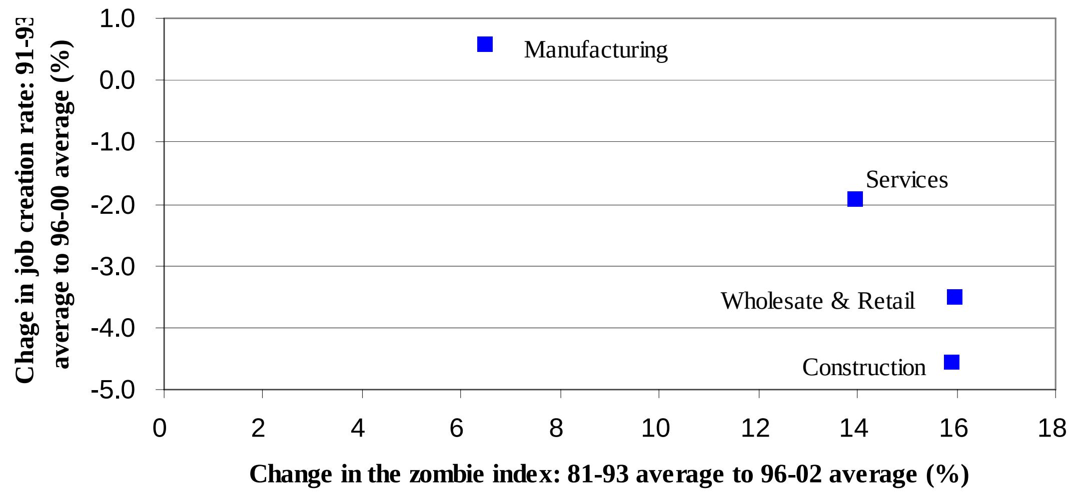


Figure 6

Zombies and TFP Growth

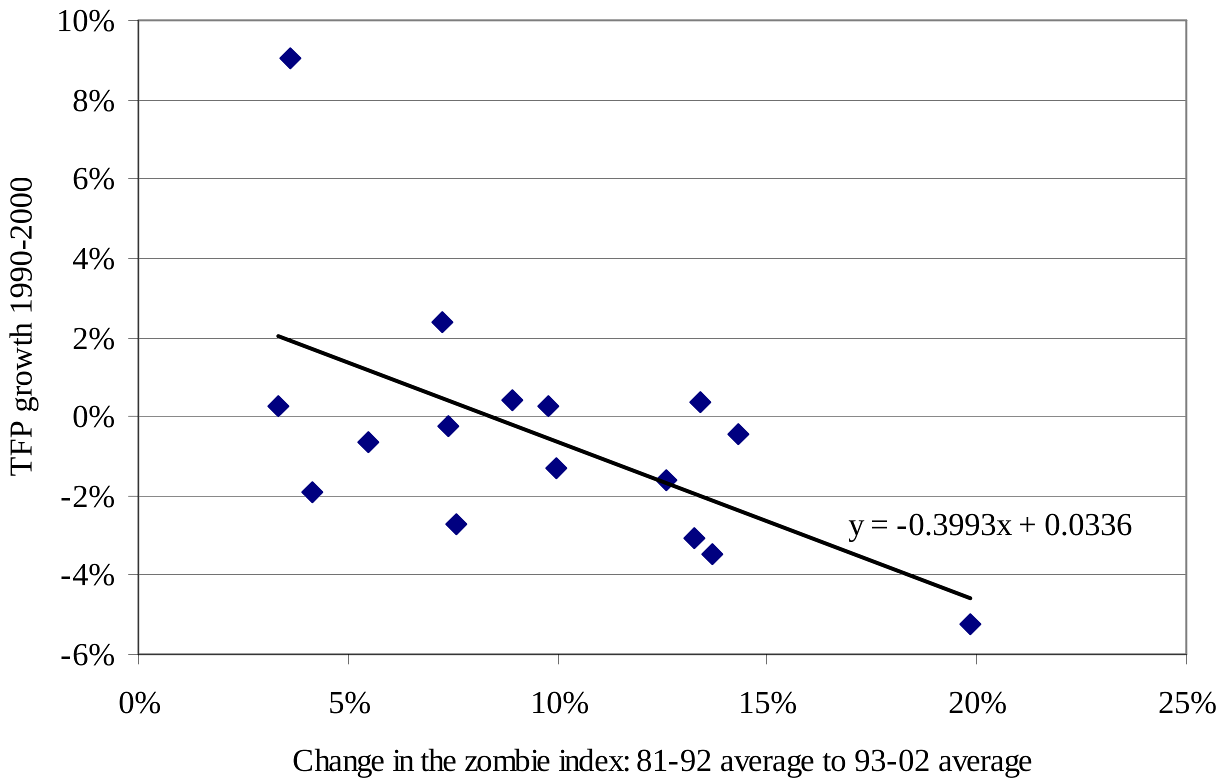


Figure 7: Marginal Effect of the Industry Zombie Percentage for Non-Zombie Firms with Different Levels of Sales Growth

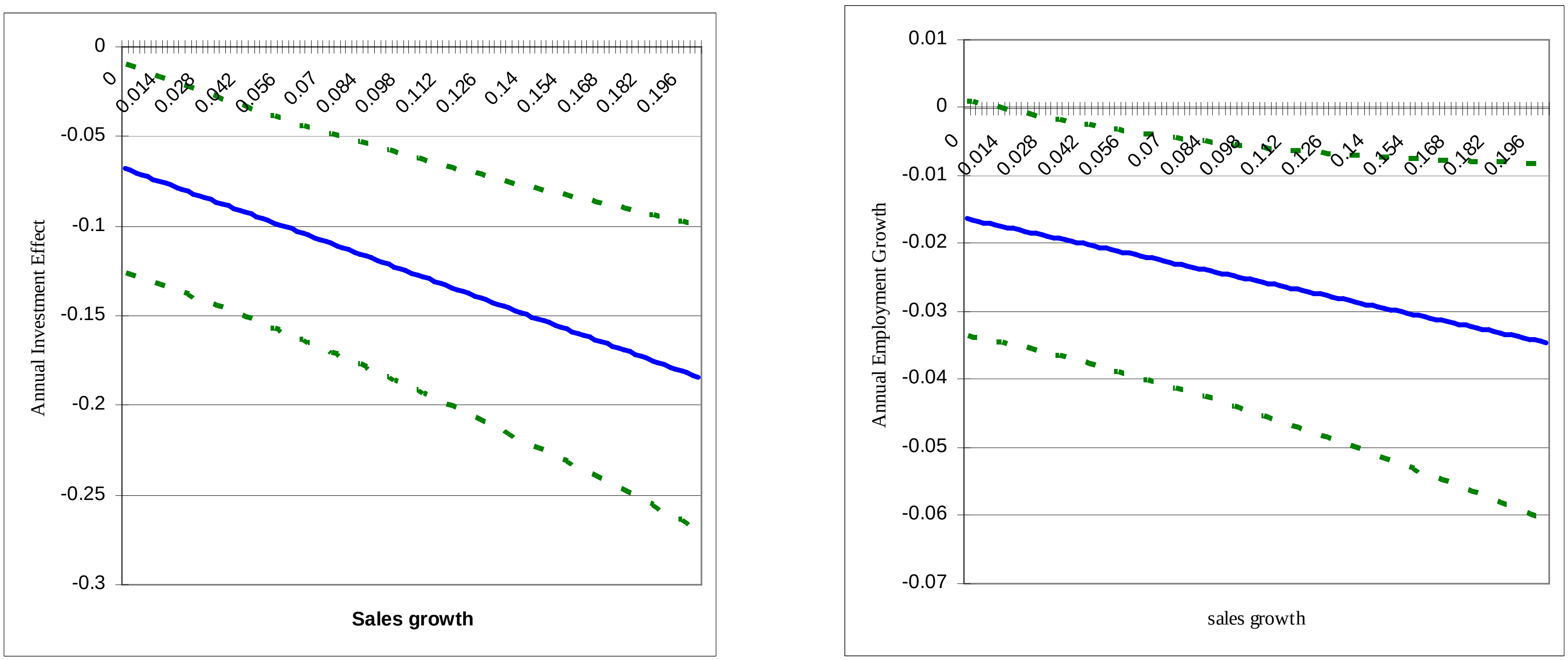


Table 1
Search Results For News Articles Regarding Restructured Companies

Total Hits for January 1990 through May 2004	1,196
Of which, related to private sector companies in Japan	1,085
Clear description of the content of "financial assistance" (excludes duplicate articles on the same case)	120
• New loans	19
• Interest concessions （金利減免）	**36**
• Purchase of new shares （新株引き受け）	29
• Debt-Equity swaps	**26**
• Debt forgiveness （債権放棄）	**44**
• Moratorium on loan principle （元本支払猶予）	**11**
• Moratorium on interest payments （利子支払猶予）	**5**

Notes: Search words: "Financial assistance" AND ("Management Reconstruction Plan" OR ("Corporation" and "Reconstruction")); actual phrases were 金融支援 AND (経営再建計画 OR (企業 AND 再建)).

Source: Nikkei Telecom 21.

Table 2
Impact of Zombie Firms on the Investment, Employment and Productivity of Non-Zombies Using Baseline Zombie Estimates

Dependent Variable	I/K	ΔLog E	Log Sales – ⅔ Log E – ⅓ Log K	I/K	ΔLog E
Sample	1993-2002	1993-2002	1993-2002	1993-2002	1993-2002
Constant	0.2390 (0.0084)	0.0137 (0.0024)	3.3842 (0.0196)	0.2465 (0.0084)	0.0162 (0.0025)
Non-Zombie Dummy	0.0256 (0.0056)	0.00109 (0.001751)	0.0139 (0.0135)	0.0241 (0.0058)	0.0009 (0.0017)
Industry Zombie %	-0.1370 (0.0376)	-0.0454 (0.0116)	-0.3418 (0.0922)	-0.0987 (0.0364)	-0.0283 (0.0108)
Non-Zombie * Industry Zombie%	-0.0885 (0.0330)	-0.0232 (0.0102)	0.2183 (0.0756)	-0.0678 (0.0297)	-0.0163 (0.0088)
Sales growth				0.1152 (0.0318)	0.1078 (0.0097)
Non-Zombie * Sales Growth				0.1436 (0.0376)	0.0160 (0.0116)
Industry Zombie% * Sales Growth				1.1002 (0.1402)	0.1674 (0.0427)
Non-Zombie * Sales Growth * Industry Zombie%				-0.5823 (0.1733)	-0.0912 (0.0535)
$\bar{R}^2$	0.0537	0.0895	0.3599	0.1083	0.1700

The sample consists of between 1,844 and 2,506 publicly traded firms (depending on the year). Each regression is estimated after trimming the top and bottom 2.5% of observations (based on the dependent variable). White (1980) standard errors are reported in parentheses under each coefficient estimate. Industry and year dummies are also included in each regression. Any firm with actual interest payments below the hypothetical minimum is considered a zombie and any firm where this is not true is a non-zombie (d_1=d_2=0 in equation (1)). Two digit industry classifications are used throughout. The industry percentages for zombies are based on the share of total industry assets residing in zombie firms. Sales growth is the log difference of each firm's sales. I/K is the ratio of investment in depreciable assets to beginning of period stock of depreciable assets (measured at book value). E is the total number of full time employees. K is the book value of depreciable assets.

Table 3
Cumulative Impact of Zombie Firms on Non-Zombies

A. Cumulative investment losses (1993-2002) of the median non-zombie firm in the high zombies industries

Industry	Wholesale	Retail	Construction	Real Estate	Services
Actual Average I/K: 1993-2002	0.1184	0.1871	0.1373	0.0920	0.2215
Cumulative Lost I/K Case 1	0.1206	0.0525	0.0833	0.0793	0.0842
Cumulative Lost I/K Case 2	0.0963	0.0399	0.0503	0.1117	0.1408

"Actual Average I/K: 1993-2002" shows the actual average investment rate (I/K) of the median non-zombie firm in the industry for 1993-2002. "Cumulative Lost I/K Case 1" shows the total amount of investment (I/K) of the typical non-zombie that was depressed during the period compared with the hypothetical case where the asset weighted zombie index had stayed at its average level for 1981-1992. "Cumulative Lost I/K Case 2" shows the total amount of investment (I/K) of the typical non-zombie that was depressed during the period compared with the hypothetical case where the asset weighted zombie index of the industry was the same as that of manufacturing in each year from 1993 to 2002. The coefficient estimates from the regression in the column 2 of Table 2 were used for the calculation.

B. Cumulative employment change (1993-2002) of the median non-zombie firm in the high zombies industries

Industry	Wholesale	Retail	Construction	Real Estate	Services
Average Actual Employment growth: 1993-2002	-0.0136	0.0015	-0.0043	0.0062	0.0134
Cumulative lost employment -- Case 1	0.0381	0.0190	0.0285	0.0301	0.0381
Cumulative lost employment -- Case 2	0.0303	0.0144	0.0172	0.0427	0.0641

"Average Actual Employment Growth: 1993-2002" shows the actual average annual rate of change in the employment at the median non-zombie in the industry for 1993-2002. "Cumulative lost employment Case 1" shows the total rate of new hiring at the typical non-zombie that was depressed during this period compared with the hypothetical case where the asset weighted zombie index had stayed at its average level for 1981-1992. "Cumulative lost employment Case 2" shows the total rate of new hiring at the typical non-zombie that was depressed during the period compared with the hypothetical case where the asset weighted zombie index of the industry was the same as that of manufacturing in each year from 1993 to 2002. The coefficient estimates from the regression in the column 3 of Table 2 were used for the calculation.

Table A-1
Impact of Zombie Firms on the Investment, Employment and Productivity of Non-Zombies
Using Fuzzy Zombie Indices

Dependent Variable	I/K		ΔLog E		Log Sales – ⅔ Log E – ⅓ Log K		I/K		ΔLog E	
$\{d_1, d_2\}$ (in basis points) in eq. (1)	{0, 50}	{-25, 75}	{0, 50}	{-25, 75}	{0, 50}	{-25, 75}	{0, 50}	{-25, 75}	{0, 50}	{-25, 75}
Constant	0.2460 (0.0086)	0.2488 (0.0090)	0.0148 (0.0024)	0.0157 (0.0026)	3.3745 (0.0200)	3.3816 (0.0210)	0.2527 (0.0085)	0.2549 (0.0088)	0.0174 (0.0024)	0.0181 (0.0025)
Non-Zombie Dummy	0.0304 (0.0061)	0.0323 (0.0068)	0.0026 (0.0019)	0.0029 (0.0021)	0.0343 (0.0144)	0.0284 (0.0162)	0.0286 (0.0060)	0.0305 (0.0067)	0.0017 (0.0018)	0.0019 (0.0020)
Industry Zombie %	-0.2016 (0.0335)	-0.2295 (0.0368)	-0.0555 (0.0100)	-0.0616 (0.0111)	-0.2504 (0.0797)	-0.3044 (0.0875)	-0.1632 (0.0327)	-0.1879 (0.0359)	-0.0417 (0.0097)	-0.0477 (0.0107)
Non-Zombie * Industry Zombie%	-0.0572 (0.0264)	-0.0583 (0.0294)	-0.0161 (0.0080)	-0.0177 (0.0090)	0.1114 (0.0615)	0.1584 (0.0684)	-0.0465 (0.0250)	-0.0478 (0.0279)	-0.0082 (0.0077)	-0.0093 (0.0086)
Sales growth							0.1137 (0.0492)	0.1074 (0.0535)	0.1083 (0.0204)	0.1113 (0.0219)
Non-Zombie * Sales Growth							0.1516 (0.0616)	0.1564 (0.0684)	0.0108 (0.0252)	0.0029 (0.0277)
Industry Zombie% * Sales Growth							0.7937 (0.2166)	0.7960 (0.2373)	0.1459 (0.0869)	0.1432 (0.0951)
Non-Zombie * Sales Growth * Industry Zombie%							-0.3982 (0.2639)	-0.3840 (0.2948)	-0.0880 (0.1207)	-0.0709 (0.1350)
$\bar{R}^2$	0.0556	0.0559	0.0897	0.0898	0.3631	0.3630	0.1096	0.1097	0.1705	0.1707

The sample consists of between 1,844 and 2,506 publicly traded firms (depending on the year). Each regression is estimated after trimming the top and bottom 2.5% of observations (based on the dependent variable). White (1980) standard errors are reported in parentheses under each coefficient estimate. Industry and year dummies are also included in each regression. The zombie probabilities are calculated as described in the text using equation (1). Two digit industry classifications are used throughout. The industry percentages for zombies are based on the share of total industry assets residing in zombie firms. Sales growth is the log difference of each firm's sales. I/K is the ratio of investment in depreciable assets to beginning of period stock of depreciable assets (measured at book value). E is the total number of full time employees. K is the book value of depreciable assets. Sample period is 1993 to 2002.

Table A-2
Impact of Zombie Firms on the Investment, Employment and Productivity of Non-Zombies
Excluding observations with the interest rate gap close to zero

Dependent Variable	I/K		ΔLog E		Log Sales – ⅔ Log E – ⅓ Log K		I/K		ΔLog E	
Range of excluded obs (in basis points)	[0, 50]	[-25, 75]	[0, 50]	[-25, 75]	[0, 50]	[-25, 75]	[0, 50]	[-25, 75]	[0, 50]	[-25, 75]
Constant	0.2331 (0.0088)	0.2377 (0.0099)	0.0129 (0.0026)	0.0133 (0.0029)	3.3776 (0.0210)	3.3789 (0.0231)	0.2413 (0.0087)	0.2455 (0.0097)	0.0153 (0.0025)	0.0152 (0.0028)
Non-Zombie Dummy	0.0293 (0.0059)	0.0251 (0.0070)	0.0019 (0.0018)	0.0018 (0.0021)	0.0613 (0.0143)	0.0468 (0.0164)	0.0267 (0.0058)	0.0223 (0.0069)	0.0013 (0.0018)	0.0013 (0.0021)
Industry Zombie %	-0.0972 (0.0390)	-0.1111 (0.0469)	-0.0318 (0.0124)	-0.0262 (0.0145)	-0.2056 (0.0989)	-0.3388 (0.1121)	-0.0611 (0.0382)	-0.0732 (0.0462)	-0.0143 (0.0120)	-0.0077 (0.0141)
Non-Zombie * Industry Zombie%	-0.1274 (0.0356)	-0.1087 (0.0415)	-0.0374 (0.0110)	-0.0383 (0.0127)	-0.0615 (0.0828)	0.0432 (0.0934)	-0.0958 (0.0340)	-0.0728 (0.0400)	-0.0262 (0.0107)	-0.0256 (0.0124)
Sales growth							0.1191 (0.0494)	0.1712 (0.0580)	0.1086 (0.0216)	0.1421 (0.0211)
Non-Zombie * Sales Growth							0.1852 (0.0621)	0.1329 (0.0717)	0.0107 (0.0249)	-0.0260 (0.0256)
Industry Zombie% * Sales Growth							1.1047 (0.3073)	0.9219 (0.3565)	0.1717 (0.1189)	0.0519 (0.1257)
Non-Zombie * Sales Growth * Industry Zombie%							-0.6630 (0.3729)	-0.3733 (0.4317)	-0.0225 (0.1500)	0.1147 (0.1636)
$\bar{R}^2$	0.0502	0.0457	0.0858	0.0792	0.3652	0.3595	0.1099	0.1086	0.1681	0.1637

The sample consists of between 1,844 and 2,506 publicly traded firms (depending on the year). Each regression is estimated after trimming the top and bottom 2.5% of observations (based on the dependent variable). White (1980) standard errors are reported in parentheses under each coefficient estimate. Industry and year dummies are also included in each regression. Any firm with actual interest payments below the hypothetical minimum is considered a zombie and any firm where this is not true is a non-zombie (d_1=d_2=0 in equation (1)). Two digit industry classifications are used throughout. The industry percentages for zombies are based on the share of total industry assets residing in zombie firms. Sales growth is the log difference of each firm's sales. I/K is the ratio of investment in depreciable assets to beginning of period stock of depreciable assets (measured at book value). E is the total number of full time employees. K is the book value of depreciable assets. Sample period is 1993 to 2002.

Table A-3
Impact of Zombie Firms on the Investment, Employment and Productivity of Non-Zombies Using Estimated Maturity Structure for Long-term Borrowings and Bonds

Dependent Variable	I/K	ΔLog E	Log Sales – ⅔ Log E – ⅓ Log K	I/K	ΔLog E
Constant	0.2496 (0.0090)	0.0169 (0.0026)	3.3919 (0.0210)	0.2528 (0.0088)	0.0180 (0.0026)
Non-Zombie Dummy	0.0125 (0.0062)	-0.0007 (0.0021)	0.0133 (0.0147)	0.0144 (0.0060)	0.0005 (0.0019)
Industry Zombie %	-0.0668 (0.0520)	-0.0388 (0.0163)	-0.3601 (0.1190)	-0.0168 (0.0493)	-0.0224 (0.0162)
Non-Zombie * Industry Zombie%	-0.0867 (0.0505)	-0.0321 (0.0155)	0.2285 (0.1122)	-0.0784 (0.0473)	-0.0288 (0.0150)
Sales growth				0.1952 (0.0561)	0.1316 (0.0214)
Non-Zombie * Sales Growth				0.0382 (0.0630)	-0.0132 (0.0248)
Industry Zombie% * Sales Growth				0.6669 (0.4490)	-0.0068 (0.1458)
Non-Zombie * Sales Growth * Industry Zombie%				0.4628 (0.4983)	0.2068 (0.2086)
$\overline{R}^2$	0.0521	0.0897	0.3614	0.1075	0.1704

The sample consists of between 1,844 and 2,506 publicly traded firms (depending on the year). Each regression is estimated after trimming the top and bottom 2.5% of observations (based on the dependent variable). White (1980) standard errors are reported in parentheses under each coefficient estimate. Industry and year dummies are also included in each regression. Any firm with actual interest payments below the hypothetical minimum is considered a zombie and any firm where this is not true is a non-zombie (d_1=d_2=0 in equation (1)). Two digit industry classifications are used throughout. The industry percentages for zombies are based on the share of total industry assets residing in zombie firms. Sales growth is the log difference of each firm's sales. I/K is the ratio of investment in depreciable assets to beginning of period stock of depreciable assets (measured at book value). E is the total number of full time employees. K is the book value of depreciable assets.

Table A-4
Impact of Zombie Firms on the Investment, Employment and Productivity of Non-Zombies Assuming Zero for the Minimum Required Interest Rate on Bonds

Dependent Variable	I/K	ΔLog E	Log Sales – ⅔ Log E – ⅓ Log K	I/K	ΔLog E
Constant	0.2382 (0.0083)	0.0131 (0.0024)	3.3834 (0.0195)	0.2464 (0.0082)	0.0158 (0.0023)
Non-Zombie Dummy	0.0237 (0.0056)	0.0007 (0.0017)	0.0129 (0.0133)	0.0223 (0.0055)	0.0006 (0.0017)
Industry Zombie %	-0.1879 (0.0394)	-0.0533 (0.0123)	-0.3915 (0.0941)	-0.1452 (0.0384)	-0.0338 (0.0120)
Non-Zombie * Industry Zombie%	-0.0793 (0.0336)	-0.0213 (0.0104)	0.2283 (0.0764)	-0.0575 (0.0320)	-0.0145 (0.0101)
Sales growth				0.1240 (0.0495)	0.1104 (0.0214)
Non-Zombie * Sales Growth				0.1394 (0.0593)	0.0144 (0.0239)
Industry Zombie% * Sales Growth				1.0730 (0.3132)	0.1561 (0.1191)
Non-Zombie * Sales Growth * IndustryZombie%				-0.5706 (0.1154)	-0.0835 (0.1489)
$\bar{R}^2$	0.0543	0.0896	0.3599	0.1084	0.1699

The sample consists of between 1,844 and 2,506 publicly traded firms (depending on the year). Each regression is estimated after trimming the top and bottom 2.5% of observations (based on the dependent variable). White (1980) standard errors are reported in parentheses under each coefficient estimate. Industry and year dummies are also included in each regression. Any firm with actual interest payments below the hypothetical minimum is considered a zombie and any firm where this is not true is a non-zombie (d_1=d_2=0 in equation (1)). Two digit industry classifications are used throughout. The industry percentages for zombies are based on the share of total industry assets residing in zombie firms. Sales growth is the log difference of each firm's sales. I/K is the ratio of investment in depreciable assets to beginning of period stock of depreciable assets (measured at book value). E is the total number of full time employees. K is the book value of depreciable assets.

THE JOURNAL OF FINANCE • VOL. LVIII, NO. 1 • FEB. 2003

Financial Distress and Bank Lending Relationships

SANDEEP DAHIYA, ANTHONY SAUNDERS, and ANAND SRINIVASAN*

ABSTRACT

We use a unique data set of bank loans to examine the wealth effects on lead lending banks when their borrowers suffer financial distress. We find a significant negative announcement return for the lead lending bank when a major corporate borrower announces default or bankruptcy. Banks with higher exposure to the distressed firm have larger negative announcement-period returns. The existence of a past lending relationship with the distressed firm results in larger wealth declines for the bank shareholders. Finally, financial distress also has a significant negative effect on borrower's returns.

RECENT BANKING LITERATURE has focused increased attention on the costs and benefits of banking relationships (see Boot (2000)). In particular, existing empirical work has been aimed primarily at establishing and estimating the value of the relationships that borrowers have established with their bankers. Leading examples include James (1987; excess returns to borrowers on the announcement of new bank loans), Peterson and Rajan (1994; the role of the length of a relationship in determining the availability and pricing of bank credit), Puri (1996) and Gande et al. (1997; better pricing of newly issued securities when the underwriter also has a lending relationship). More recently, Dahiya et al. (2002) examine the value loss when bank relationships weaken by examining loan sale announcements. Thus, there is considerable evidence of value creation for the borrower, on initiation or renewal of a banking relationship. However, there is a paucity of studies that attempt to measure the value of such relationships for lending banks. Peterson and Rajan (1995) and Berlin and Mester (1998) suggest that banks have incentives to smooth out the interest charged on loans if they have repeated transactions with a borrower over a long period of time. This suggests that banks find it valuable to invest in and maintain long-term customer relationships. Nevertheless, the costs of such relationships to a bank are often ignored. In this paper, we attempt to fill this gap in the literature by examining the impact of a borrower's distress on its lead bank. The financial distress of a borrower should

*Dahiya is with Georgetown University, Saunders is with New York University, and Srinivasan is with the University of Georgia. This paper has benefited greatly from suggestions and comments from an anonymous referee as well as from Rick Green (the editor), Yakov Amihud, Ned Elton, Loretta Mester, David Yermack, Marc Zenner, and Kose John. We would like to thank Edward Altman, Mark Carey, and Edith Hotchkiss for providing the data used in this study.

reduce the value of any banking relationship. Specifically, we analyze the bank's share price reaction when one of a bank's corporate borrowers enters financial distress as reflected by a bond default and/or bankruptcy.

There are strong arguments for treating a borrower's distress announcement as a "no news" (or low cost) event for a bank. First, prudent banking norms limit the losses that a bank might suffer if any single borrower is unable to repay its debt, and typically a bank loan is secured and is senior debt.[1] Thus, the recovery rate on defaulted debt is likely to be fairly high for bank loans.[2] Second, banks are considered insiders with significant informational advantages. This implies that banks are likely to be better informed about the financial status of their borrowers and thus will be able to take steps to reduce their loan exposures before the news of a borrower's distress becomes public information. Last, in many instances, news of a borrower's distress is preceded by other public announcements such as a decline in its earnings, a cut in its dividends, and so forth, which may diminish the informational content of news regarding financial distress.

These effects have to be weighed against the hypothesis that a borrower's financial distress announcement is an "adverse news" event that has a negative impact on a bank's share price. First, there is a direct effect on the bank due to the expected losses because of the borrower's distress. This effect should be related to the exposure of the bank to the borrower. Second, there may also be indirect effects of the borrower's distress on the bank's stock price. Such indirect effects may arise from many sources. For example, multiplier or contagion effects may exist if the distress of one borrower is correlated across an industry (or region). That is, a firm's distress may convey information about an increased likelihood of distress of other borrowers in the same industry to which the bank may be exposed. In addition, the news of a corporate borrower's distress may be construed as a signal of poor loan initiation and management skills, with an accompanied loss of a bank's reputational values.[3] Last, bank regulators scrutinize banks' books to ensure that they meet the banking requirements of regulatory capital. A default or bankruptcy event is likely to increase this scrutiny and, therefore, will act as an additional "regulatory" tax on the bank.[4]

[1] Regulation also restricts loan exposures to a borrower to a maximum of 15 percent of the capital of the bank.

[2] Weiss (1990) studies 37 publicly traded firms and finds that secured creditors' claims are paid in full in most cases. To the extent that bank loans are secured, this suggests a fairly high recovery rate. Franks and Torous (1994), using a sample from 1983 to 1989, find that bank loans have recovery rates of about 85 percent. However, more recent evidence by Gupton, Gates, and Carey (2000) suggests a recovery rate of between 50 to 65 percent for bank loss in the event of default.

[3] The loss of reputational value may also be reflected in the unwillingness of other banks to enter into new loan syndications arranged by the lead bank. For example, Smith (1992) reports that Salomon Brothers lost over a third of it market value because of the treasury auction scandal in 1991. This loss of value, which amounted to over $1.5 billion, was well above the fines and other costs arising from expected legal and regulatory sanctions.

[4] Also, Benveniste, Singh, and Wilhelm (1993) find that the bankruptcy of Drexel Burnham resulted in positive returns for rival banks likely to benefit from this event.

The central hypothesis explored in this paper is whether the news of a corporate borrower's distress has a material economic impact on its lead bank. A group of studies has examined the impact on the lending banks' share price of the announcement of debt moratoriums by sovereign borrowers (see, e.g., Lamy, Marr, and Thompson (1986), Smirlock and Kaufold (1987), Grammatikos and Saunders (1990), and Musumeci and Sinkey (1990)). However, the evidence from these studies is mixed, with the majority of them finding negative reactions that are heterogeneous across banks.

Kracaw and Zenner (1996) examine bank share price reactions to nine highly leveraged firms that became financially distressed. They find a negative share price reaction for these banks, but one that was not statistically significant. However, these findings were for a very small sample of firms involved in highly leveraged transactions such as leveraged buyouts (LBOs) or recapitalizations. In contrast, our sample consists of a much larger number of firms (many of which are publicly traded) that faced financial distress and/or bankruptcy over a relatively lengthy sample period.

The rest of the paper proceeds as follows. Section I describes the sample selection and the data collection procedure. Section II describes the methodology. The empirical results and their interpretation are presented in Section III. We conclude with a brief summary of our main findings in Section IV.

I. Sample Selection and Data Collection

We define a firm to be financially distressed if it has insufficient cash flows to meet the payments on its debt.[5] This paper examines two types of financial distress announcements: (1) the default on a firm's public debt, and (2) the filing by a firm for bankruptcy protection under Chapter 11. Specifically, in all cases, we use a single date for default for each company, and this date is the first date of default of publicly traded debt as listed in the database on defaulted bonds produced by Altman.[6]

Gilson et al. (1990) (henceforth, GJL) and others use a broader set of events to define distress. In particular, while defaults constitute over 50 percent of the first event of distress in GJL, they also use reports of restructuring of debt, where the firm's creditors suffer some impairment on their claim due to an exchange of debt into equity or convertible securities, reduction of interest payments or

[5] Gilson, John, and Lang (1990), Wruck (1990), Franks and Torous (1994), Tashijian, Lease, and McConnell (1996), and Andrade and Kaplan (1998) study the effect of distress on the borrowing firm.

[6] We used data from "The Defaulted Bond Master Database" produced and updated by Altman (see, e.g., Altman and Kishore (1996)). This database includes defaults on medium- and long-term publicly traded bonds and debentures and does not include short-term debt or commercial paper defaults. This database covers over 1,000 defaults beginning in 1971. Default is defined as the issuer's inability to meet an interest or principal payment on the company's medium-term or long-term public debt, and the date on which this occurs is defined as the date of default.

principal, or extension of maturity. Such restructuring is termed as distressed restructuring. Towards the end of this paper, we study the effect of such distress events on our results. For the most part, our sample consists of distress events that occur subsequent to the sample period in GJL. Also, the borrower does not have to be a publicly traded firm to enter into our sample; rather, the only requirement is that the borrower have a bank loan outstanding.[7]

The impact of each of these announcements on a borrower's lead bank is estimated by calculating the abnormal returns for the bank's shareholders around the date of the relevant announcement. The study analyzes 71 cases of default and 101 cases of bankruptcy filings for a 10-year period, 1987 to 1996. We study the two events separately by constructing a sample of lead banks that had loans outstanding to firms on the date of their default and another sample of lead banks that had loans outstanding to firms on the date of their Chapter 11 filings. We also create a subsample of firms that are common to both samples. For this latter subsample, we first examine the impact of default and then that of subsequent bankruptcy. This allows us to control for the partial anticipation of bankruptcy induced by an earlier bond default. We repeat the same test for the subsample of distressed firms that either did not have any public debt outstanding or for whom the default and the bankruptcy events occurred simultaneously. This allows us to examine the informational content of a bankruptcy announcement without contamination due to a prior signal such as default.

To construct the sample of distressed firm announcement dates and the lending relationships of these firms, the following data were employed: (1) a comprehensive list of firms that defaulted on their public debt and the date of the default over the period 1987 to 1996, (2) a list of firms that filed for Chapter 11 and their filing date over the period 1987 to 1996, (3) details of bank loans made to distressed firms that were outstanding at the time of their default and/or bankruptcy.

The primary source for the list of defaults was an updated version of the Altman database used in Altman and Kishore (1996), and the primary source for the bankruptcies was Hotchkiss (1995).[8] These lists were cross-checked and supplemented with information from a variety of other sources. These included *The Bankruptcy Almanac*, published by New Generation Research, and various news sources, such as the Dow Jones News Retrieval Services and the Lexis-Nexis bankruptcy library.

We used data from the Loan Pricing Corporation Database (LPC)[9] to get details of loan transactions and the nature of the relationship between distressed firms and their banks. The LPC database contains detailed transaction-level information related to loan amount, start and expiration dates, terms and purpose of the loans, the name of the lead bank(s), and the syndicate's size. These details are especially advantageous in examining the impact of a borrower's distress on its lead bank. First, it allows for the identification of the lead bank(s) of a

[7] Tashijian et al. (1996) also do not require presence of public stock or private debt. However, they focus only on prepackaged bankruptcies.

[8] We thank Edith Hotchkiss and Ed Altman for sharing their data.

[9] We thank Mark Carey for providing help with the LPC data.

distressed firm. Second, the details of the loan transactions provide a rich cross section of loan attributes. Last, the start and the maturity dates allow us to determine whether a loan was outstanding at the time of distress.

The sample selection procedure was as follows: The names of the firms that defaulted on their bonds or filed for bankruptcy were hand matched with the list of loan borrowers in the LPC database. This allowed us to determine for which of these firms loan data were available.[10] This step yielded a list of 971 loan transactions involving borrowers that subsequently defaulted or filed for bankruptcy. This sample was further narrowed down to include only those transactions which were entered into before the date of the distress announcement and that had a contractual maturity date later than the date of the distress announcement. This step insured that we only included those transactions that could reasonably be assumed to be outstanding at the onset of financial distress. Using the default date as the date for distress, we obtained a sample of 174 transactions that could be assumed to be outstanding on the date of default. A similar procedure using the date of Chapter 11 filing as the distress date yielded a sample of 272 transactions. Next, we located the lead bank for these transactions by looking and searching for the words "arranger," "administrative agent," "agent," or "lead bank" in the lender role definition in the LPC database.[11] Finally, we used the names of all the commercial banks listed on CRSP in conjunction with the list of lead banks for the sample of firms that defaulted or filed for Chapter 11.[12] This step eliminated the transactions that had a foreign bank or a nonbank finance company as the lead lender.

The final sample consists of 123 transactions involving 71 borrowers that subsequently defaulted on their public debt and 174 transactions involving 101 borrowers that subsequently filed for bankruptcy. The 71 announcements of default represent 99 announcement events for the lead banks, while the 101 announcements of bankruptcy represent 130 announcement events. The higher number of bank events compared to the number of firms in distress occurs because some of the firms had multiple lead banks.

Sources of data on bank and borrower characteristics for the distressed borrowers' sample included the LPC database, Moody's manuals, and BANK COMPUSTAT. The size of the loan transaction, the purpose and terms of the deal, and the syndicate size were taken from the LPC database. The balance sheet data for

[10] The LPC database lists each credit facility as a separate record field. Thus, a single borrower may have multiple credit facilities from the same bank, or a single credit facility that is syndicated among multiple banks, or multiple facilities syndicated among multiple banks.

[11] The LPC database lists the role of the lead syndicate member as arrangers, co-arrangers, lead bank, agents, co-agents, documentation/credit agent, or lead manager. We looked for any of these role levels within a syndicate to assign the lead bank role. All the credit facilities included in the sample had a clear lead bank, as all syndicate members other than the lead bank were defined as participants in the credit facility.

[12] A number of loans were made by the subsidiaries of the main bank holding company. For this study, we use the stock price reaction of the bank holding company. Also, some of the banks merged or were taken over after the loan was made, but before the company filed for bankruptcy. In those instances, the announcement effect of bankruptcy is measured on the merged/acquirer bank.

lead banks, where available, was obtained from BANK COMPUSTAT data tapes. Where the information was not available, it was supplemented by information taken from Moody's bank manuals. The information on bank holding companies was obtained from Moody's bank manuals.

In Panel A of Table I we provide the distribution of the sample by the year of the bankruptcy and default announcement. Most of the financial distress announcements are clustered in the years 1990 to 1993. This is in line with the timing of the economic recession of the 1990 to 1991 period, when more cases of financial distress are to be expected. More than 70 percent of the defaults and bankruptcies occur in the 1990 to 1993 period.

In panel B, we document the characteristics of the lender banks. There are 22 different banks for the sample of 71 firms that defaulted on their public debt (36 banks for 101 firms that filed for Chapter 11). We also report the mean ratio of total transaction size to the equity of the lead bank in the year before the date of distress. This ratio is 19.02 percent (median 6.08 percent) for the banks of defaulting firms (mean 12.1 percent, median 2.7 percent for the banks of firms filing for Chapter 11). The absolute level of this ratio should be interpreted with care as it represents the ratio of the aggregate transaction size to the lead bank's equity. This is not the bank's actual exposure, as that would depend on the share of the

Table I

Descriptive Statistics of the Sample

This table presents the descriptive statistics of the credit transactions and the lead banks of the borrowers that experienced financial distress over the period 1987 to 1996. The sample is categorized by the year of onset of financial distress; i.e., default on public debt or filing for chapter 11 (in Panel A), the number of lead banks involved (in Panel B), the size and type of transaction (in Panel C), and the specific structure of the credit transaction (in Panel D).

Panel A: Distribution of the sample by year of announcement

Year of financial distress	Number of defaults	Number of bankruptcies
1987	2	1
1988	1	1
1989	3	7
1990	16	16
1991	27	32
1992	11	22
1993	3	11
1994	2	4
1995	6	6
1996	0	1
Total (1987–1996)	71	101

Panel B: Lender characteristics

	Default	Chapter 11 filing
Number of bank announcement events	99	130
Number of different banks involved	22	36
Average transaction size to bank equity ratio (median)	19.02% (6.8%)	12.10% (2.7%)
Average loan loss reserve ratio (median)	2.3% (2.05%)	2.08% (2.05%)

Table I.—*continued*

Panel C: Mean (median) statistics for the credit transactions

Transaction Type	Default			Chapter 11		
	No. of Trans-actions	Trans-action size ($ mm)	Syndicate size	No. of Trans-actions	Trans-action size (mm)	Syndicate size
Debt repayment/ consolidation	26	208 (145)	10.4 (5.5)	45	156 (80)	9.9 (6)
General corporate purposes	21	236 (100)	9.2 (6)	36	162 (50)	7.4 (4)
Working capital	25	161 (110)	8.2 (7)	31	128 (78)	7 (2)
Takeover acquisition	17	962 (200)	22.1 (7)	18	912 (212)	20.8 (8)
Leveraged buyout	16	616 (327)	25.9(17)	14	380 (357)	20 (8)
Recapitalization	10	919 (370)	33 (19.5)	10	816 (287)	29 (15)
Others	8	167 (112)	10.8 (5)	20	101 (53)	5 (1.5)
TOTAL	123	415 (195)	15.2 (7)	174	280 (260)	11.4 (4)

Panel D: Additional information on the structure of credit transactions

	Default	Chapter 11 filing
	Mean (Median)	Mean (Median)
Premium over LIBOR (n = 88; 96)	2.05% (2.50%)	1.91% (2.00%)
Premium over US prime (n = 110; 151)	1.24% (1.50%)	1.17% (1.37%)
Commitment fee (n = 74; 88)	0.46% (0.50%)	0.42% (0.50%)
Up-front fee (n = 59; 81)	1.21% (1.15%)	1.02% (0.87%)

transaction size retained by the lead bank, since most of these transactions are syndicated among other participating banks and nonbank finance companies. Also, each transaction is made up of multiple facilities, not all of which may be fully drawn-down. While this ratio is not an exact estimate of the bank's exposure to the distressed firm, it is a proxy for the upper limit of a bank's exposure.[13]

In Panel C, we provide the descriptive statistics at the aggregate loan transaction level. Our full sample consists of 123 transactions involving firms that defaulted and 174 transactions involving firms that filed for Chapter 11. All of these loan transactions had a contractual life that overlapped the date of default/bankruptcy by the borrowing firm. The mean (median) transaction size of $415 million ($195 million) for the default sample is larger than the mean (median) transaction size of $280 million ($260 million) for the bankruptcy sample. We also partition our sample by the stated purpose of the transaction. The large proportion of the lending is for the purposes of LBOs, recapitalizations, takeovers, and working capital. Although takeover/acquisition, leveraged buyouts, and

[13] Indeed, a bank faces reputational losses, in addition to the size of loan retained, should a syndicate led by that bank make losses. The mean ratio of loan loss reserves to the total assets of the lead banks for the year before the year of distress is 2.3 percent (median 2.05 percent) for the sample of lead banks of defaulting firms, and 2.1 percent (median 2.05 percent) for the banks of bankrupt firms.

recapitalizations account for less than one-fourth of the number of transactions, they account for over 60 percent of the total dollar value. These transactions are also fairly large. For example, in the case of the sample of firms that declared bankruptcy, the mean transaction size for takeover/acquisition is $912 million (median $212 million), for LBOs $380 million ($357 million), and for recapitalizations $816 million ($287 million). In comparable studies of LBOs and recapitalizations, Kaplan (1989) has a mean (median) transaction size of $524 million ($254 million), and Muscarella and Vetsuypens (1990) have a mean (median) value of $250 million ($105 million). Relatively large syndicates finance these transactions. The mean size of the syndicate involved in a transaction to finance takeover/acquisitions is 20.8 (median 8), while that for LBOs is 20 (median 8), and for recapitalizations 29 (median 15). The transactions providing the day-to-day regular financing for working capital and general corporate purposes account for over 45 percent of the transactions by number, but constitute less than 25 percent of the total dollar value. This is reflected in the mean transaction size for working capital, $128 million (median $ 78 million), and for general corporate purposes, $157 million ($50 million). The mean syndicate size for working capital is 7 (median 2), and 7.1 (median 4) for general corporate purposes.

Additional information about the structure of the sample of transactions is presented in Panel D of Table II. For the sample of firms filing for Chapter 11, on

Table II

Cumulative Abnormal Returns for the Lead Banks on the Announcement of Financial Distress by Their Borrowers

Cumulative abnormal return (CARs) for the lead banks of the firms facing financial distress over the period 1987 to 1996. Panel A describes the share performance of lead bank(s) around the date of default by the borrower, and Panel B describes the same around the date of chapter 11 filing. The sample of firms that defaulted on their public debt and the date of default are obtained from the Altman database of defaulted bonds, and the sample of firms that filed for chapter 11 is compiled from multiple sources including Hotchkiss (1995), DJNR, the *Bankruptcy Almanac*, and Lexis bankruptcy library. CARs are calculated using the Center for Research in Security Prices (CRSP) database.

Event Window	ACAR	t-statistic	Median CAR	Wilcoxon z-statistic
Panel A: Abnormal returns for the lead banks when borrowers default on public debt ($N = 99$)				
11-day window [– 8, 2]	– 3.766%	– 5.327***	– 2.547%	– 3.886***
7-day window [– 4, 2]	– 1.922%	– 3.445***	– 1.158%	– 3.084***
5-day window [– 2, 2]	– 1.750%	– 3.769***	– 1.522%	– 3.192***
3-day window [– 1, 1]	– 0.857%	– 2.459**	– 0.719%	– 2.513**
Panel B: Abnormal returns for the lead banks when borrowers file for chapter 11 ($N = 130$)				
11-day window [– 8, 2]	– 1.832%	– 3.303***	– 1.305%	– 2.165**
7-day window [– 4, 2]	– 1.089%	– 2.437**	– 0.686%	– 2.003**
5-day window [– 2, 2]	– 0.911%	– 2.131**	– 0.350%	– 2.268**
3-day window [– 1, 1]	– 0.492%	– 1.367	– 0.386%	– 1.391

***Significant at the one percent level
**Significant at the five percent level
*Significant at the 10 percent level

average, the borrower paid a premium of 1.89 percent (median 2 percent) over LIBOR and 1.16 percent (median 1.25 percent) over the U.S. prime rate. Commitment fees on the unused portion of the lending facility are 0.43 percent and the up-front fees are 1 percent. The statistics for the sample of defaulting firms are similar. The loan rates are lower for our sample compared to those reported by Kracaw and Zenner (1996) for their sample of nine highly leveraged transactions.

II. Test Methodology

The basic null hypothesis is:
H_0: The announcement of a firm's financial distress is a "no news" event for the firm's lead bank(s).
The alternative hypothesis is:
H_1: The announcement of a firm's financial distress is an "adverse news" event for the firm's lead bank(s).

A simple way to examine the reaction of a bank's stock price to the announcement of a borrower's financial distress is to employ a standard event study methodology to study the movements in the bank's stock return around the date of the announcement of financial distress by the borrower. However, the use of the default announcement or Chapter 11 filing as the study event poses some problems. These announcements are usually preceded by many other announcements and news stories that foreshadow the subsequent announcement of default and/or bankruptcy. Thus, the traditional narrow event window of two or three days is unlikely to capture the entire stock price reaction of a bank lender. A typical chronology of the release of various distress-related announcements is illustrated by Figure 1.

The release of relevant information prior to the actual event is well illustrated by the chronology of news items that appeared before the default and bankruptcy of Columbia Gas Systems, which suffered a long period of financial deterioration before finally filing for bankruptcy.

- June 20, 1991—Columbia Gas Systems suspends dividends and calls for renegotiations with its gas suppliers. The firm said that potential losses on existing contracts exceed $1.1 billion. (*Wall Street Journal*)
- June 21, 1991—Columbia Gas Systems defaulted on $15 million of commercial paper and other short-term notes. (*Wall Street Journal*)

1	2	3	4
Dividend cuts, Earnings decline	Default on debt, Credit rating change	Chapter 11 filing	Plan of reorganization is confirmed

Figure 1. Typical chronology of distress-related news information. The figure provides a typical time-line illustration of how different classes of distress-related news stories are announced.

- June 25, 1991—Columbia Gas Systems defaulted on an additional \$10 million of short term notes. (*Wall Street Journal*)
- July 9, 1991—Columbia Gas Systems defaulted on \$14 million of commercial paper. (*Wall Street Journal*)
- July 22, 1991—Columbia Gas Systems defaulted on \$15 million of short-term debt. (*Wall Street Journal*)
- August 1, 1991—Columbia Gas Systems files for bankruptcy. (*Wall Street Journal*).

The long drawn-out nature of the distress process makes the use of the standard two- or three-day event window unsuitable for this study.[14] To fully capture the impact of the deterioration in the bank–borrower relationship more usefully, we use four different event windows: 11 days, 7 days, 5 days, and the traditional 3 days, to measure the market's reaction to the news of a borrower's default and bankruptcy on its lead bank. (We also examined event windows of 15 and 21 days; the results are essentially identical and are not reported).

We calculate the announcement abnormal returns for banks using the market model methodology as detailed in Mikkelson and Partch (1986) and James (1987). The parameters of the market model are estimated by regressing the firm's common stock returns for the period 200 days before the event date, to 50 days before the event date on the rate of return on CRSP's dividend inclusive, equal-weighted index for NYSE/AMEX/Nasdaq stocks. The abnormal return is computed as the difference between the observed return and the estimated return from the market model. Cumulative abnormal returns (CAR) are the sum of abnormal returns for the days in the relevant event window. Tests of significance are based on standardized abnormal returns and CARs. In the last section of the paper, we also examine the effect of distress on the borrower using an event study in a manner similar to GJL and others. We use the equally weighted index as the benchmark and a similar estimation period for the borrowing firms.

III. Empirical Results

A. Stock Price Response to the News of Financial Distress

If the announcement of its borrowers' financial distress is a no news event for the lead bank, we do not expect to find any abnormal movement in the stock price of the bank around the date of the announcement. However, if the distress is an adverse news event, then negative wealth effects for the lead bank's shareholders are expected. The results presented below are largely consistent with the adverse news hypothesis.

[14] The date of default in the Altman database for Columbia Gas Systems was July 31, 1991 (see footnote 6). This date also corresponds to the date on which the company's senior and subordinated debt were downgraded to the default rating status of D by S&P.

Specifically, Table II, Panel A, presents the average stock price response of the lead banks to the announcements of public debt defaults of their borrowers. For the 11-day period starting 8 days before the news of default and lasting until 2 days after, the average cumulative abnormal return (ACAR) is -3.77 percent, which is significant at the 1 percent level (t-statistic = -5.39). $-$ Narrowing the event window to 7 days and further to 5 days leaves the results unchanged. For the traditional event window of 3 days (-1, 0, +1), the ACAR is -0.86 percent, which is still significant at the 5 percent level (t-statistic = -2.45). Thus, the news of borrowers' defaulting on their public debt is received as significant adverse news by the shareholders of the lead lending banks.

In Table II, Panel B, we repeat the event study for a different sample of firms for whom the onset of financial distress is proxied by the date of their filing for Chapter 11 bankruptcy. The direction of the results is similar to the default sample—again, the announcement has a negative effect on the lead bank's share price. However, the scale of the stock price reaction is much lower on the news of bankruptcy than on bond default. The 11-day ACAR is -1.83 percent, which is significant at the 1 percent level (t-statistic = -3.30). This is roughly half of the size of the price reaction that banks suffer on the news of default. (The results are similar for 7-day and 5-day windows). For the 3-day window, the ACAR is negative but statistically insignificant.[15]

Overall, the tests for the aggregate sample of banks provide strong evidence for the adverse news hypothesis especially on the announcement of bond defaults. However, there is considerable variation in the size of credit transactions and the size of the lending banks. Although the details of individual bank exposures are not public knowledge, an informationally efficient market would react more severely to the distress announcement of a borrower in which the lead bank had a relatively high exposure compared to a bank with a relatively low exposure. We examine the impact of exposure levels on the magnitude of the lead bank's stock price reaction next.

B. Stock Price Response of High versus Low Exposure Banks

If the individual share of each bank in each loan transaction were known, we would be able to determine the exact dollar amount that the lead bank has directly exposed to the distressed firm. Unfortunately, we only have data on the aggregate size of each loan transaction. To differentiate the banks with high exposure from those with low exposure, we calculate the following

[15] We also performed a nonparametric Wilcoxon signed rank test for differences between the abnormal returns. This test has two principal advantages over parametric tests like the t-test and the z-test: (1) It uses only the rank of the returns ordered in terms of the magnitude and sign and does not use the magnitude of the return; therefore, it is robust to the presence of outliers; and (2) it does not make assumptions about the distribution of abnormal returns. If the two sets of data are identically distributed, then the sum of ranks of the two sets should be close to each other. The difference in the sum of ranks can be used to test for differences in the sample mean of the two data sets. In virtually all cases (See Table II, column 5), the direction and the level of statistical significance of the results of this test are consistent with those obtained in the t-tests.

exposure ratio for each bank j to a given borrowing firm k:

$$Exp_j = \frac{\sum_{i=1}^{T} Transacamount_{ij}^{k}}{Bankcap_j}, \tag{1}$$

where for each borrower k, $Transacamount_{ij}$ is the dollar amount of transaction i; for which bank j was the lead bank, T is the total number of loan transactions that lead bank j has outstanding at the time of distress to the borrower, and $Bankcap_j$ is the capital of the bank j as reported for the year before the year of borrower's distress.

Thus, Exp_j provides a proxy for the bank's exposure to the distressed borrower. We divide the default sample into the two subsamples based on this ratio: banks with exposure ratios higher than the median of 6.84 percent and banks with exposure levels lower than the median. This is repeated for the bankruptcy sample (median exposure ratio 2.78 percent). In Table III, we report the results of the event study for these subsamples. We find that the price reaction is much more negative and significant for the subsample of highly exposed banks. This holds true for both the news of a bond default as well as for bankruptcy. Panel A compares the ACAR for the high exposure banks and the low exposure banks around the date of default by the bank's borrower. The 11-day ACAR is -5.48 percent for the high exposure banks, which is approximately twice as large as the ACAR of -2.09 percent for the low exposure banks. The ACAR is negative and statistically significant for the high exposure banks across all event windows, yet while it is also negative for the low exposure banks, it is not statistically significant. We also test whether the difference between the ACAR for the two groups is statistically significant. The last column in panel A reports the t-statistics for the difference between high exposure and low exposure banks. The differences between the two groups are significant for all event windows.

Panel B reports the same results for the price reaction around the date of Chapter 11 filings. The results are similar to the ones reported in Panel A. The 11-day ACAR for the high exposure banks is -2.95 percent (t-statistic $= -3.49$). Low exposure banks, on the other hand, have an ACAR of -0.72 percent (t-statistic $= -1.17$). Varying the length of the event window to 7, 5 or 3 days leaves the results largely unchanged, as the high exposure banks suffer a price reaction much larger than that for the banks with low exposure. However, the statistical significance of the difference between the two groups is much weaker. As reported in the last column, the difference between the two groups is only significant for the 11-day window.

Finally, some of the firms in our bankruptcy sample are included in the sample of defaulting firms. This may diminish the true effect of the bankruptcy announcement for these firms, as the news of their bankruptcy may have already been anticipated by the news of their bond default. Thus, the chronological order of various distress announcements may have an important bearing on how the market reacts to news regarding distressed borrowers. We investigate this in the next section.

Table III

Cumulative Abnormal Returns for the Higher and Lower Exposure Lead Banks on Announcement of Financial Distress by Their Borrowers

Average Cumulative Abnormal Return (ACARs) for the lead bank(s) of the firms facing financial distress over the period 1987 to 1996. The figures in parentheses below each ACAR represent the ± statistic that tests if the given ACAR is significantly different from zero. Panel A compares the share performance of higher exposure lead banks with that of lower exposure lead bank(s) around the date of default by the borrower, and Panel B describes the same around the date of Chapter 11 filing by the borrower. The exposure is defined as the ratio of aggregate transaction size divided by the lead bank's equity as reported for the latest year before borrowers' financial distress. The sample of firms that defaulted on their public debt and the date of default are obtained from the Altman database of defaulted bonds, and the sample of firms that filed for Chapter 11 is compiled from multiple sources including Hotchkiss (1995), DJNR, the *Bankruptcy Almanac*, and Lexis bankruptcy library. CARs are calculated using the Center for Research in Security Prices (CRSP) database.

Event Window	ACAR for higher exposure banks	ACAR for lower exposure banks	*t*-statistic for Difference
Panel A: ACAR for the lead banks around the date of default by the banks' borrowers			
11-day window [− 8, 2]	− 5.485%	− 2.081%	− 2.035**
	(− 5.517)***	(− 2.034**)	
7-day window [− 4, 2]	− 3.212%	− 0.792%	− 2.374**
	(− 4.125)***	(− 0.763)	
5-day window [− 2, 2]	− 3.086%	− 0.441%	− 2.784***
	(− 4.559)***	(− 0.790)	
3-day window [− 1, 1]	− 1.829%	0.098%	− 2.697***
	(− 3.692)***	(− 0.202)	
Panel B: ACAR for the lead banks around the date of Chapter 11 filing by the banks' borrowers			
11-day window [− 8, 2]	− 2.946%	− 0.719%	− 1.701*
	(− 3.499)***	(− 1.173)	
7-day window [− 4, 2]	− 1.855%	− 0.321%	− 1.624
	(− 2.909)***	(− 0.536)	
5-day window [− 2, 2]	− 1.268%	− 0.554%	− 0.979
	(− 2.127)**	(− 0.886)	
3-day window [− 1, 1]	− 0.894%	− 0.095%	− 1.418
	(− 1.739)*	(− 0.201)	

***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

C. Reaction to the News of Default Prior to Bankruptcy

The sample of borrowing firms filing for bankruptcy can be divided into two subsamples depending on whether or not there was a bond default prior to the bankruptcy. We construct a subsample of 33 firms that defaulted on their public debt at least seven days prior to filing for bankruptcy. The remaining 68 firms either did not have any public debt outstanding or their default and bankruptcy announcements occurred on the same day.

For the subsample of 33 firms, the announcement of a Chapter 11 filing would be partially anticipated because of their prior default on public debt. Thus, the expectation of bankruptcy would already have been incorporated (in part) in the bank's share price by the time bankruptcy was actually announced. For the subsample of 68 firms, however, the news of the bankruptcy would still have significant informational content. This is supported by the results reported in Table IV. Panel A presents the ACAR of the banks for the first subsample (33 firms) on the announcement of default, while Panel B reports the ACAR results for the

Table IV

Cumulative Abnormal Returns for the Lead Banks for the Sub-sample of Borrowers Filing for Bankruptcy

Average Cumulative Abnormal Return (ACARs) for the lead bank(s) of the firms that defaulted on their bonds at least seven days prior to filing for bankruptcy over the period 1987 to 1996. Panels A and B describe the share performance of the lead banks of the firms that defaulted on their public debt at least seven days before filing for bankruptcy. Panel A presents the ACAR around the date of default and Panel B describes the same around the date of bankruptcy by the borrowers. Panel C describes the performance of lead banks for the subsample of bankrupt firms that either did not have public debt outstanding or the default and bankruptcy announcement was made simultaneously. The sample of firms that defaulted on their public debt and the date of default are obtained from the Altman database of defaulted bonds, and the sample of firms that filed for chapter 11 is compiled from multiple sources including Hotchkiss (1995), DJNR, the *Bankruptcy Almanac*, and Lexis bankruptcy library. CARs are calculated using the Center for Research in Security Prices (CRSP) database.

Event window	ACAR	t-statistic
Panel A: ACAR for the lead banks on the date of default for their borrower firms that subsequently filed for bankruptcy ($N = 51$)		
11-day window [−8, 2]	−2.812%	−3.011***
7-day window [−4, 2]	−1.929%	−2.787***
5-day window [−2, 2]	−1.291%	−2.256**
3-day window [−1, 1]	−0.669 %	−1.561
Panel B: ACAR for the lead banks on the date of bankruptcy by the firms in Panel A sample ($N = 51$)		
11-day window [−8, 2]	−0.187%	0.358
7-day window [−4, 2]	−0.163%	0.083
5-day window [−2, 2]	−0.321%	0.008
3-day window [−1, 1]	0.0409%	0.611
Panel C: ACAR for the lead banks on the date of bankruptcy by the firms that had no public debt or the default and bankruptcy occurred simultaneously ($N = 79$)		
11-day window [−8, 2]	−2.894%	−4.267***
7-day window [−4, 2]	−1.686%	−3.193***
5-day window [−2, 2]	−1.293 %	−2.739***
3-day window [−1, 1]	−0.836%	−2.243**

***Significant at the one percent level
**Significant at the five percent level
*Significant at the 10 percent level

same 33 firms in the event of their subsequent bankruptcy filing. As reported in Panel A, the announcement returns are -2.81 percent for the 11-day window on the news of the bond default, which is significant at the one percent level (t-statistic $= -3.01$). The results are directionally similar for the 7-, 5-, and 3-day windows. However, when these same firms declare bankruptcy (Panel B), the ACAR for the lead banks is only -0.19 percent for the 11 days around the date of their Chapter 11 filing, failing to reject the null hypothesis (t-statistic = 0.358). The results are similar for event windows of shorter length. These results imply that prior news of a bond default significantly reduces the informational content (for banks) of subsequent bankruptcy announcements by their borrowers.

In Panel C, we report the results for the subsample of bankrupt firms that either did not have any public debt or for whom default and the bankruptcy occurred simultaneously. For these firms, there is no default signal of distress prior to their declaration of bankruptcy. In the absence of any prior default news, we find a bankruptcy announcement return for banks of -2.89 percent for the 11-day window, which is significant at the one percent level (t-statistic $= -4.27$). The results are robust to different lengths of the event window, which continue to be negative and significant. Thus, when there is no prior bond default, the bankruptcy announcement of a firm has a significantly negative impact on the market value of its lead lending bank.[16]

D. Multivariate Tests

Our results so far show that the shareholders of a lead bank suffer a wealth decline when there is unanticipated news of financial distress by their major borrowers. In this section, we seek to confirm our univariate findings and to investigate other factors that may potentially affect bank abnormal returns around distress announcements. Specifically, there is considerable variation in the characteristics of the banks and the loans in our sample, as well as in macroeconomic conditions, which may have had an impact on how the market reacted to news of a borrower's distress. To examine the impact of these factors on announcement period returns for banks lending to distressed firms, we estimate a regression model that takes the following form:

$$\begin{aligned}\mathrm{CAR_j} = \beta_0 &+ \beta_1\,\mathrm{RELATIONSHIP}_j + \beta_2\,\mathrm{EXPOSURE_j} + \beta_2\,\mathrm{RECESSION}_j \\ &+ \Sigma\beta_k\,\mathrm{CNTRLVAR}_{jk},\end{aligned} \tag{2}$$

[16] It should be noted that once news of a prior default becomes public, the reputational losses as well as the possibility of additional regulatory scrutiny are realized. Such losses could very well incorporate the possibility of a bankruptcy and subsequent losses in that process. Even though the recovery rates are high, there is still a large amount of uncertainty in the amount to be recovered. Using recent evidence, Gupton et al. (2000) suggest that mean bank loan value in default is 69.5 percent for senior secured debt and 52.1 percent for senior unsecured debt. However, loss given default values have a large variance with the lowest 10th percentiles of recoveries at 39.2 percent for senior secured debt and 5.8 percent for senior unsecured debt.

where:

- CAR_j is the dependent variable is the 11-day cumulative abnormal return for bank j around the date of bankruptcy (or bond default) by the bank's borrower. (For the borrowers that are common to the default and bankruptcy samples, the earlier of the two events is used.)
- $RELATIONSHIP_j$ is a dummy variable that takes on the value one if bank j had been the lead bank in lending (making previous loans) to the distressed borrower before its default/bankruptcy. This variable captures the existence of a prior bank-borrower relationship.
- $EXPOSURE_j$ is a measure of the exposure of bank j to the distressed borrower as defined by equation (1) earlier.
- $RECESSION_j$ is a dummy variable that takes the value of one if the distress occurs between the dates of July 1, 1990, and March 31, 1991 (the peak-to-trough business contraction dates as defined by the National Bureau for Economic Research.)
- $CNTRLVAR_{jk}$ is a set of control variables for loan and bank characteristics. These include the following.
  - LOAN LOSS RESERVE is the loan loss reserve of the bank divided by the bank capital in the year prior to the distress date.
  - LOAN LOSS RESERVE DIFFERENCE is the difference of the bank's loan loss reserves for the year of the distress date and its loan loss reserves for the year prior to the distress date divided by the bank capital in the year prior to distress.
  - BANKSIZE is the natural log of the total assets of the bank as reported for the year prior to the date of distress.
  - LBO is a dummy variable that takes the value of one if the loan purpose was for a leveraged buyout.
  - CREDIT SPREAD is the spread of the loan over LIBOR at the time of loan origination.
  - MULTIPLE BANK DUMMY is set to one when the borrower involved in the distress event has multiple lead banks.
  - PRIOR DISTRESS DUMMY is set to one when the first event of distress was not a default on public debt or a bankruptcy (e.g., it was a debt restructuring) and we can identify the exact date.[17]
  - INDUSTRY: A set of dummy variables to control for the borrower's industry.

E. Regression Results

Higher loan exposures should put a lending bank at risk of losing a greater proportion of its capital base, and thus risk insolvency or closure if the borrower is unable to repay its loans. This implies a negative relationship between announcement period returns and the bank's degree of exposure. This is indeed the case in Models 1 and 3 in Table VA.

[17] See the earlier papers of Gilson et al. (1990) for a definition of distress that includes debt restructuring.

Table VA

Regressions Relating the Cumulative Abnormal Return for the Lead Bank around the Date of Financial Distress to the Lender and Borrower Characteristics

The OLS regression of cumulative abnormal returns (CAR) for the 11-day window around the dates of distress (bankruptcy in case there is no prior news of distress, default if it occurs before bankruptcy). The independent variables include the following: RELATIONSHIP is a dummy variable that takes the value one if the bank provided credit to the firm in the past. EXPOSURE is a ratio of the aggregate sum of all credit facilities extended to the distressed firm by the lead bank to the total equity of the lead bank as reported for the year before the year of distress. RECESSION DUMMY is a dummy variable that is one if the date of distress is between July 1, 1990, and March 31, 1991. LBO is a dummy variable that takes the value one if the loan transaction was for the purpose of leveraged buyout. In addition to the variables reported, the regression also includes industry dummies based on the one-digit SIC code of the borrower. Numbers in the parentheses are standard errors.

Variable	(1)	(2)	(3)
INTERCEPT	0.0139	0.0171	0.0181
	(0.0191)	(0.0172)	(0.0172)
EXPOSURE	− 0.0434*		− 0.0295*
	(0.0149)		(0.0165)
RELATIONSHIP		− 0.0379***	− 0.0343***
		(0.0114)	(0.0121)
RECESSION DUMMY	− 0.0556***	− 0.0521***	− 0.0517***
	(0.0179)	(0.0167)	(0.0166)
LBO DUMMY	0.0256	0.0269	0.0266
	(0.0215)	(0.0184)	(0.0184)
N	156	156	156
Adj. *R* Sq.	0.10	0.13	0.14

***Significant at the one percent level
**Significant at the five percent level
*Significant at the 10 percent level

As discussed in the introduction, if the bank had been involved with the borrower in a lending relationship prior to its distress, the relationship is likely to have been of value and the dissolution of such a relationship is likely to be costly to the bank.[18]

Thus, we expect that banks, which had a prior lending relationship with a distressed borrower, will be more adversely impacted by the onset of distress. To measure this effect, an indicator variable, RELATIONSHIP, is used, which equals one if the bank has been involved in a lending relationship with the distressed firm prior to its distress or bankruptcy.

The regression results reported in Table VA provide strong support for the argument that prior relationships are valuable. Distress of a borrower with a past

[18] Slovin, Shushka, and Polonchek (1993) document significant value loss for the *borrowers* of a *bank* (Continental Illinois) when it was facing distress. Others such as Lummer and McConnell (1989) find that loan renewals result in positive abnormal returns for borrowers. Our study examines this impact in the other direction, that is, the effect of a borrower's distress on the lender.

Table VB

Regressions Relating the Cumulative Abnormal Return for the Lead Bank around the Date of Financial Distress to the Lender and Borrower Characteristics

The OLS regression of cumulative abnormal returns (CAR) for the 11-day window around the dates of distress (bankruptcy in case there is no prior news of distress, default if it occurs before bankruptcy). See Table V A or Section III Subsection D for the definitions of the exposure, relationship, recession, and LBO variables. BANKSIZE is the natural log of the total assets of the lead bank as reported for the year prior to the year in which distress occurs. The LOAN LOSS RESERVE is the loan loss reserve of the bank in the year prior to the date of distress divided by the bank's capital in that year. The difference between the loan loss reserve of the bank at the end of the year of the first distress event (default or bankruptcy) and the loan loss reserve in the year before the distress event divided by the bank's capital in the year before distress is the LOAN LOSS RESERVE DIFFERENCE. EXP $\times$ LLRD is the interaction of the exposure and the loan loss difference variables. CREDIT SPREAD is the spread of the loan over LIBOR at the time the loan was issued. The PRIOR DISTRESS DUMMY takes a value of one if the first distress event was not a default or a bankruptcy and the company experienced financial distress as defined in Gilson, John, and Lang (1990). The MULTIPLE BANK DUMMY takes a value of one when the borrower has multiple lead banks. In addition to the variables reported, the regression also includes industry dummies based on the one-digit SIC code of the borrower. Numbers in the parentheses are standard errors.

Variable	(1)	(2)	(3)	(4)
INTERCEPT	0.0072	0.0034	0.0184	– 0.0236
	(0.0217)	(0.0222)	(0.0174)	(0.0408)
RELATIONSHIP	– 0.0340***	– 0.0336***	– 0.0337***	– 0.0308***
	(0.012)	(0.0121)	(0.0121)	(0.0123)
RECESSION DUMMY	– 0.0554***	– 0.0564	– 0.0509***	– 0.0722***
	(0.0166)	(0.0165)	(0.0166)	(0.0273)
LBO DUMMY	0.0253	0.0253	0.0272	0.0415*
	(0.0182)	(0.0182)	(0.0184)	(0.0203)
LOAN LOSS RESERVE	0.0278	0.0379		
	(0.0303)	(0.0326)		
LOAN LOSS RESERVE DIFFERENCE (LLRD)		0.0394 (0.0449)		
LLRD $\times$ EXP			0.0797*	
			(0.0442)	
BANKSIZE				0.0003
				(0.0002)
CREDIT SPREAD				– 0.0131
				(0.0117)
PRIOR DISTRESS DUMMY				– 0.0055
				(0.0122)
MULTIPLE BANK DUMMY				– 0.0056
				(0.0124)
N	156	156	156	153
Adj. R Sq.	0.14	0.14	0.14	0.14

***Significant at the one percent level
**Significant at the five percent level
*Significant at the 10 percent level

relationship with the bank is relatively more costly for the bank. Specifically, the coefficient for the prior relationship variable (RELATIONSHIP) is negative and significant at the one percent level as reported in models 2–3, Table VA.

One possible concern about these relationship results is that they are driven by a few transactions that have large negative returns. Out of the total sample of 156 transactions, 68 transactions involved borrowers and banks that had prior relationships. These 68 transactions involved 62 different firms and 13 different banks. Moreover, the ACAR (recorded over 11 days) for this subsample with prior relationships was −5.04 percent and that for the subsample without prior relationships was −0.81 percent. The difference in these abnormal returns was significantly different from zero at the one percent level of significance. Thus, both the univariate and multivariate tests suggest that the existence of a prior relationship is important in impacting the scale of the valuation effect on a bank with loans outstanding to a distressed borrower.[19]

The RECESSION dummy variable controls for the different macroeconomic conditions prevailing at the time of the announcement of distress. Our sample period, 1987 to 1996, includes the 1990 to 1991 economic recession. The negative and significant coefficient for the recession dummy variable is consistent with the view that the news of financial distress has a larger negative impact on the lending banks during a period of economic contraction.[20]

Next, in Table VB, we investigate the effect of other bank and firm specific variables on the abnormal returns of the leading bank. In particular, our previous univariate and multivariate results suggest that one source of the loss to banks is the direct loss arising from the size of their loan exposure to the distressed borrower. To investigate this effect further, we use data on loan loss reserves in the regression. These are reserves that banks are required to set aside against expected or anticipated future losses on their loan portfolio. Thus, anticipation of a distress event should result in an increase of the banks' loan loss reserve ratio in the period prior to distress. If a bank has built up sufficient reserves, it is less likely to fail as a result of borrower defaults.[21] As such, we should find that banks that have built up loan loss reserves prior to default should be less negatively impacted by news of distress events.[22] Model 1 in Table VB tests this possible relationship. As can be seen, the loan loss reserve variable is positive but insignificant. In Model 2, we test if additions to the reserve have any impact on the abnormal returns. Other authors such as Grammatikos and Saunders (1990)

[19] The studies by Lummer and McConnell (1989) used only existence of a prior relationship in evaluating the stock price reaction. Other authors such as Peterson and Rajan (1994) and Berger and Udell (1995) also use the duration of the relationship as a measure of the strength of the relationship. Unfortunately, the LPC data is censored, starting only from the beginning of 1987 and, therefore, we cannot measure the duration of the relationship.

[20] One possible reason for this is that recovery rates, including the value of collateral such as real estate, are likely to be lower in recessions and contractions.

[21] Loan loss reserves can be viewed as a first line of defense against losses (i.e., expected losses), while capital reserves can be viewed as the second line of defense, that is, against unexpected losses.

[22] We thank the referee for suggesting this.

found that additions to the loan loss reserve by banks were viewed favorably by the market. This variable is also found to be positive but insignificant. Lastly, we interact the difference in loan loss reserves with the exposure variable. We find that this variable has a positive and significant effect, suggesting that banks that have larger exposures and increase their loan loss reserves in anticipation of financial distress of one of their borrowers tend to have less negative abnormal returns than banks that do not.[23]

We also controlled for the LBO loans and the borrower's industry by the inclusion of a set of dummy variables. Except for the LBO dummy, which was significantly positive in one model, none of these control variables had regression coefficients that were statistically significant in Tables VA and VB.

F. Robustness Checks

As mentioned earlier, our definition of distress is somewhat more restrictive than that used in the study by GJL. To test the robustness of our findings, we collected data on the first date of distressed restructuring of the borrowers in our sample (see Section I for a definition of a distressed restructuring). We go back two years prior to the first date of distress as defined in this study (which is a public default or bankruptcy) to find a date when (or if) the borrower attempted to restructure its debt. We used the Lexis-Nexis libraries on public news and bankruptcy for this search. We focused only on those restructuring events where it was clear from the related news story that the attempted restructuring of debt was due to financial difficulties. In several cases, the restructuring event coincided with the default or bankruptcy date. In those cases where the company attempted to restructure its debt before default or bankruptcy (and we can identify the exact date when this happened), we set the PRIOR DISTRESS DUMMY to 1 in the multivariate tests in Table VB, Model 4; otherwise, the dummy is given a value of zero. As can be seen, while the sign of the dummy is negative, it is not statistically significant in our sample.

We control for bank-specific characteristics with the variables BANKSIZE. On one hand, one might expect the importance of any individual corporate loan default to be relatively small, because larger banks are likely to be more diversified. On the other hand, the borrowers of larger banks tend to be bigger and more widely followed companies. The distress announcement of such a borrower may cause a larger negative reaction for a larger bank as the market revises its assessment of the quality of the bank's overall loan portfolio and efficiency of the bank as a "delegated monitor" (see, e.g., James (1987)). As can be seen from Table VB, Model 4, the bank size variable was found to be insignificant.

The coefficient on the CREDIT SPREAD variable (the spread on the loan over LIBOR at the time of loan origination) should proxy for the bank's ex ante expectation of the borrower's risk before the distress event. As can be seen, inclusion of

[23] One complicating factor is that banks appear to use the loan loss reserve not only as a fund to insulate against future losses but also to smooth earnings. See Collins, Shackelford, and Wahlen (1995), Beaver and Engel (1996), and Wall and Koch (2000) for evidence on such smoothing.

the credit spread on the loan made by the bank prior to distress had no significant explanatory effect on the size of the lending bank's CAR at the time of distress.

A potential problem in our multivariate test results is that each event of distress for a given borrower may result in multiple events in our regression if the given borrower had multiple lead banks. To control for this, we created a MULTIPLE BANK DUMMY for those distress events where the borrower had multiple lead banks (dummy equal to 1) and those where it had a single lead bank (dummy equal to 0). As can be seen from Table VB, Model 4, this dummy appears to be insignificant. Indeed, univariate tests of the difference in ACARs between these two categories (i.e., multiple bank lending versus single bank lending) indicated that these two sets of ACARs were not statistically different from each other.[24]

Finally, our findings, while supportive of the adverse information hypothesis, may be confounded, in part, by some other event that negatively affects both the value of the bank and the borrowing firm, but has nothing (directly) to do with the distress itself. A good example of such confounding events is an increase in prime lending rates (see Park, Nabar, and Saunders (1993) for evidence of the effect of prime rate changes on bank returns). To account for this, we collected data on prime rate increases that occurred during the 11-day event window around distress announcements. Excluding events contaminated by prime rate increases had no effect on our results.[25]

G. Borrower Returns

Thus far, we have examined the returns to the borrower's lead banks around the distress dates. A natural line of inquiry would be to examine the effect of distress on the borrower itself. If adverse information about the borrower is indeed the cause of the negative return experienced by the bank, the borrower should also experience negative abnormal returns during the event window period. We estimate borrower returns using the methods suggested by GJL, employing an estimation period from 250 days before the announcement date to 50 days before the announcement date. Since a number of the companies in our sample were private and many had been delisted before the bankruptcy event, the number of borrowers for which these returns are available is smaller than the size of our full sample. Out of the 33 borrowers in the default subsample, 23 subsequently filed for bankruptcy. Thus, our borrower sample consists of many firms that failed to restructure their debt and therefore filed for bankruptcy.

[24] It should be noted that the returns of the two banks, although based on the same distress event, are unlikely to be identical. First, the banks' loan exposures are likely to be different. Second, their sizes will generally be different. Last, each individual bank may or may not have a prior relationship with the borrower. Therefore, the same default or bankruptcy event can have different wealth implications for the different lead banks.

[25] We thank the referee for pointing this out. We focus only on events where there is an increase in prime rates, as a decrease in prime rates is unlikely to cause negative returns to banks or their borrowers. These three events were excluded in the estimation of Model 4 in Table VB.

These results (Table VI, Panels A and B) suggest that the borrowers experienced large negative abnormal returns around the events of both default and bankruptcy. In Panel C, we present evidence on abnormal returns from debt restructuring, in those cases when the first distress event was not a default or a bankruptcy. As can be seen, these events also have significantly negative effects on borrower returns. Our results for the borrower returns (whether because of default or bankruptcy or distressed restructuring) are similar in magnitude to those obtained by GJL. For example, GJL find two-day returns of −16.7 percent for firms that file for bankruptcy. We find a one-day return of −13.4 percent for firms that file for bankruptcy. Similarly, GJL find that firms that ultimately file for Chapter 11 have a negative return of −6.3 percent at the first announcement

Table VI

Cumulative Abnormal Returns for Borrowers on their Announcement of Financial Distress

Cumulative abnormal return (CARs) for the borrower firms facing financial distress over the period 1987 to 1996. Panel A describes the share performance of firm around the date of its default on public debt, and Panel B describes the same around the date of chapter 11 filing. Panel C shows the abnormal return around the date of the first restructuring, provided this restructuring was not a default or a bankruptcy. Average cumulative abnormal return (ACAR) and *t*-statistics are calculated using methods similar to Gilson, John, and Lang (1990) and Tashijian, Lease, and McConnell (1996). The sample of firms that defaulted on their public debt and the date of default is from Altman (1996), and the sample of firms that filed for chapter 11 is compiled from multiple sources including Hotchkiss (1995), DJNR, the *Bankruptcy Almanac*, and Lexis bankruptcy library. CARs are calculated using the Center for Research in Security Prices (CRSP) database.

Event window	Number of firms	ACAR	*t*-statistic
Panel A: Abnormal returns for the borrowers when they default on public debt			
11-day window [−8, 2]	33	−10.209%	−5.674***
7-day window [−4, 2]	33	−7.378%	−4.840***
5-day window [−2, 2]	33	−4.910%	−3.786***
1-day window [day 0]	33	−2.420%	−5.306***
Panel B: Abnormal returns for the borrowers when they file for chapter 11			
11-day window [−8, 2]	43	−22.462%	−9.488***
7-day window [−4, 2]	43	−21.739%	−9.034***
5-day window [−2, 2]	43	−18.568%	−8.714***
1-day window [day 0]	43	−13.357%	−17.236***
Panel C: Abnormal returns for the borrowers when they announce distressed restructuring before the default or bankruptcy			
11-day window [−8, 2]	31	−6.683%	−4.690***
7-day window [−4, 2]	31	−8.145%	−6.512***
5-day window [−2, 2]	31	−8.696%	−7.958***
1-day window [day 0]	31	−5.752%	−9.362***

***Significant at the one percent level
**Significant at the five percent level
*Significant at the 10 percent level

of distress. The borrowers that announced a debt restructuring in our sample had a one-day negative return of -5.75 percent.[26, 27]

More importantly, we also find that the borrowers themselves had significant negative ACARs within the same event window that their lead lender banks had negative abnormal returns. Thus, the linkage between borrower distress and the negative abnormal returns of their lead banks is made stronger. In the bank event study, we found that the default event had a strong negative effect on the bank return while the bankruptcy event (when preceded by a default) had a weaker effect. For borrowers, we find that both default and bankruptcy events (as well as distressed restructurings) result in large negative abnormal returns. This suggests that these events generally had material news effects for both the borrowing firms and the banks. Perhaps, not surprisingly, the effect on borrowing firms' stock returns is larger (in percentage terms) due to the greater loss exposure of equity holders in distressed firms. By contrast, bank stockholders hold relatively senior debt claims on the borrowing firm, and, as such, normally have priority over the borrowing firm's equity holders.

IV. Conclusion

The risk of loan default is the one of the most important risks faced by banks. While there have been studies examining the impact of sovereign loan defaults on the stock prices of lending banks, a similar exercise has not been undertaken to analyze the impact of defaults/bankruptcy announcements of corporate borrowers on lending banks. The small size of any individual corporate loan relative to the size of a bank, the relatively high recovery rate for senior secured bank loans, and the prior anticipation of a borrower's financial difficulties, aligned with the role of the bank as an insider or "delegated monitor," all imply that the news of any single corporate distress might not have a significant impact on the lending bank's share price. Alternatively, industry and geography-wide correlations among distressed firms, the loss of valuable customer relationships, and the cost of lost reputation and increased regulatory scrutiny because of a borrower's distress imply that the news of a default or a bankruptcy might have a materially adverse impact on the share price of the lead lending bank.

This paper is the first large-sample documentation of the wealth effects for lead bank shareholders when bank borrowers face financial distress. For a lead bank, the news of default of a corporate borrower is associated with an average decline of 3.8 percent in its stock returns over an 11-day period surrounding the date of default. News of a corporate bankruptcy is associated with a decline in bank stock returns of 1.8 percent over a similar 11-day window. When banks are ranked according to their exposures to distressed firms, the price decline for the low exposure banks is insignificant, while that for the high exposure banks is large and significant. Our multivariate tests also indicate that exposure of a

[26] We thank the referee for suggesting this entire section.
[27] All of these borrowers did not subsequently file for Chapter 11.

bank significantly affects the size of the (negative) abnormal returns on the announcement of distress.

We also find that prior banking relationships are valuable for lenders. On average, abnormal returns to banks, on the announcement of a borrower's financial distress, are significantly and negatively related to existence of a prior past borrowing relationship with that borrower. Finally, we find that the announcement of distress also has a significantly negative effect on borrower returns in our sample. This is consistent with the results of prior studies looking at the effect of distress on borrowers.

References

Altman, Edward, and Vellore Kishore, 1996, Almost everything you wanted to know about the recoveries on defaulted bonds, *Financial Analysts Journal*, Nov./Dec, 57–64.

Andrade, Gregor, and Steven Kaplan, 1998, How costly is financial (not economic) distress? Evidence from highly leveraged transactions that became distressed, *Journal of Finance* 53, 1443–1493.

Beaver, William H., and Ellen E. Engel, 1996, Discretionary behavior with respect to allowances for loan losses and the behavior of security prices, *Journal of Accounting and Economics* 22, 177–206.

Benveniste, Lawrence M., Manoj Singh, and William J. Wilhelm Jr., 1993, The failure of Drexel Burnham Lambert: Evidence on the implications for commercial banks, *Journal of Financial Intermediation* 3, 104–137.

Berger, Allen N., and Gregory F. Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of Business* 68, 351–381.

Berlin, Mitchell, and Loretta Mester, 1999, Deposits and relationship lending, *Review of Financial Studies* 12, 579–607.

Boot, Arnoud, 2000, Relationship banking: What do we know? *Journal of Financial Intermediation* 9, 7–25.

Collins, Julie H., Douglas A. Shackelford, and James M. Wahlen, 1995, Bank differences in the coordination of regulatory capital, earnings, and taxes, *Journal of Accounting Research* 33, 263–291.

Dahiya, Sandeep, Manju Puri, and Anthony Saunders, 2002, Bank borrowers and loan sales: New evidence on the uniqueness of bank loans, *Journal of Business*, forthcoming.

Franks, Julian R., and Walter N. Torous, 1994, A comparison of financial recontracting in distressed exchanges and Chapter 11 reorganizations, *Journal of Financial Economics* 35, 349–370.

Gonde, Amar, Manju Puri, Anthony Saunders, and Ingo Walter, 1997, Bank underwriting of debt securities: Modern Evidence, *Review of Financial Studies* 10, 1175–1202.

Gilson, Stuart, Kose John, and Larry Lang, 1990, Troubled debt restructurings: An empirical study of private reorganizations of firms in default, *Journal of Financial Economics* 27, 315–354.

Grammatikos, Theoharry, and Anthony Saunders, 1990, Addition to bank loan-loss reserves: Good news or bad news?, *Journal of Monetary Economics* 25, 289–304.

Gupton, Greg M., Daniel Gates, and Lea V. Carty, 2000, Bank loan loss given default, *Moody's Investors Service Global Credit Research* Special Comment, November.

Hotchkiss, Edith, 1995, Post-bankruptcy performance and management turnover, *Journal of Finance* 50, 3–21.

James, Christopher, 1987, Some evidence of the uniqueness of bank loans, *Journal of Financial Economics* 19, 217–235.

Kaplan, Steven, 1989, The effects of management buyouts on operating performance and value, *Journal of Financial Economics* 24, 217–254.

Kracaw, William A., and Marc Zenner, 1996, The wealth effects of bank financing announcements in highly leveraged transactions, *Journal of Finance* 51, 1931–1946.

Lamy, Robert E., Wayne Marr, and G. Rodney Thompson, 1986, The Mexican debt crisis, the IMF, and the efficiency of bank share prices, *Studies in Banking and Finance* 3, 203–217.

Lummer, S., and John McConnell, 1989, Further evidence on the bank lending process and the capital market response to bank loan agreements, *Journal of Financial Economics* 25, 99–122.

Mikkelson, Wayne, and Megan Partch, 1986, Valuation effects of securities offerings and the issuance process, *Journal of Financial Economics* 15, 31–60.

Muscarella, Chris J., and Michael R. Vetsuypens, 1990, Efficiency and organizational structure: A study of reverse LBOs, *Journal of Finance* 45, 1389–1413.

Musumeci, James J., and Joseph H. Sinkey, 1990, The international debt crisis, investor contagion, and bank security returns in 1987: The Brazilian experience, *Journal of Money, Credit and Banking* 22, 209–220.

Park, Sang Yong, Prafulla Nabar, and Anthony Saunders, 1993, Prime rate changes: Is there an advantage in being first, *Journal of Business* 66, 69–92.

Peterson, Mitchell A., and Raghuram G. Rajan, 1994, The benefits of lending relationships: Evidence from small business data, *Journal of Finance* 49, 3–37.

Peterson, Mitchell A., and Raghuram G. Rajan, 1995, The effect of credit market competition on lending relationships, *Quarterly Journal of Economics* 110, 407–443.

Puri, Manju, 1996, Commercial banks in investment banking: Conflict of interest or certification role? *Journal of Financial Economics* 40, 373–401.

Slovin, Myron B., Marie E. Sushka, and John A. Polonchek, 1993, The value of bank durability: Borrowers as the bank stakeholders, *Journal of Finance* 48, 247–266.

Smirlock, Michael, and Howard Kaufold, 1987, Bank foreign lending, mandatory disclosure rules, and the reaction of bank stock prices to the Mexican debt crisis, *Journal of Business* 60, 347–364.

Smith, Clifford W., 1992, Economics and ethics: The case of Salomon Brothers, *Journal of Applied Corporate Finance* 5 (2), 23–38.

Tashijian, Elizabeth, Ronald C. Lease, and John J. McConnell, 1996, Prepacks: An empirical analysis of prepackaged bankruptcies, *Journal of Financial Economics* 40, 135–162.

Wall, Larry D., and Timothy W. Koch, 2000, Bank loan loss accounting: A review of the theoretical and empirical evidence, *Federal Reserve Bank of Atlanta Economic Review* Second Quarter, 1–19.

Weiss, Lawrence A., 1990, Bankruptcy resolution: Direct costs and violation of priority of claims, *Journal of Financial Economics* 27, 285–314.

Wruck, Karen H., 1990, Financial distress, reorganization, and organizational efficiency, *Journal of Financial Economics* 27, 419–444.

ELSEVIER Journal of Banking & Finance 23 (1999) 1067–1093

Journal of BANKING & FINANCE

www.elsevier.com/locate/econbase

Bank interest rates and credit relationships in Italy

Claudio D'Auria, Antonella Foglia, Paolo Marullo Reedtz *

Banca d'Italia, Vigilanza Creditizia e Finanziaria, Servizio Normativa e Affari Generali, Via Nazionale 187, 00184 Rome, Italy

Received 21 February 1997; accepted 13 October 1998

Abstract

When evaluating the performance of a financial system in supporting the investment activity of the corporate sector, a distinction is usually drawn between "banking economies" and "market economies", the former being characterized by long-term relations between banks and the industrial sector. Although theoretical studies and empirical results seem to agree that lending relationships increase the available quantity of capital to firms, they have little to say on the cost of bank credit: it is not clear whether a close relationship with a main bank would allow the borrower to pay a lower interest rate or expose him to a monopolistic rent. Using a unique data-set reporting detailed information on the evolution over time of individual bank–borrower relationships in Italy, we show evidence that a main bank provides credit at a lower cost and that some competition helps to reinforce the commitment between the borrower and the bank. © 1999 Elsevier Science B.V. All rights reserved.

JEL classification: G21; G32

Keywords: Interest rates; Corporate finance; Multiple banks; Lending relationships

* Corresponding author. Tel.: +39 06 4792 4553 (or +39 06 4792 4515); fax: + 39 06 4792 4460.

PII: S 0 3 7 8 - 4 2 6 6 (9 8) 0 0 1 3 1 - 9

1. Introduction

When evaluating the performance of a financial system in supporting the investment activity of the corporate sector, a distinction is usually drawn between "banking economies" and "market economies". Mayer (1994) defines banking (market) economies as those with a small (high) proportion of quoted companies, high (low) concentration of ownership, and long-term (short-term) relations between banks and industry.

It is not easy to define the Italian financial system according to these standards. Italy's private capital market is still underdeveloped in relation to the economy's capacity for production and saving. However, for large and medium-sized firms, credit relationships are not concentrated in a small number of intermediaries, as in Germany and Japan.

One branch of the literature has focused on the problem of multibank lending versus main bank lending. With regard to the price of lending, a main bank could either provide capital at a lower cost or exploit its monopolistic position by charging higher interest rates (the so-called "hold-up" problem). By pitting banks against one another, borrowers could reduce the average cost of their debt. On the other hand, a multiplicity of lenders could result in coordination problems in monitoring and bargaining costs in bankruptcy; these costs will be embedded in the rate charged.

Previous empirical studies have examined the effect of banking relationships on the cost of credit. This paper provides a more direct test for the existence of the hold-up problem. We use a unique data-set reporting very detailed information about individual bank–borrower relationships over time in Italy.

2. Literature review

Mayer (1988) argued that countries with bank-dominated financial systems, such as Germany and Japan, have experienced higher economic growth than those with the most advanced and competitive financial markets. He explained this apparent paradox as resulting from the favorable effects of close ties between the banks and the corporate sector on the availability (and cost) of credit for the firms.

The characteristics of credit relationships have been analyzed to ascertain how they affect the conduct of both banks and borrowers. As argued by Petersen and Rajan (1994) a bilateral credit relationship may be considered as an enduring commitment between a lender and a borrower who have dealt with each other for a long time and expect to continue to do so in the future.

Diamond (1991) underlines that this type of business relationship is not confined to the granting of credit; the bank can take advantage of its monitoring activity to supply financial services at a lower cost than other banks, thereby developing further sources of information. The relationship may develop to the point of banks being represented on corporate boards and, where permitted, holding corporate equity.

With regard to the investment activity of firms, banks in Japan and Germany act as a risk sharer, willing to support borrowers over the short run in the expectation of future earnings. A long-term credit relationship provides the bank with better information about the firm's financial position and its prospects and makes it convenient to support creditworthy companies even in periods of financial distress. The mutual commitment between the bank and the firm avoids future benefits from a successful rescue to be competed away (Mayer, 1988; Sheard, 1989; Hoshi et al., 1990; Calomiris, 1993; Aoki and Patrick, 1994).

The development of long-term credit relationships calls for favorable conditions with respect to the firm's ownership. A long-term commitment can be undertaken only when ownership is highly concentrated; large numbers of dispersed owners cannot commit themselves to maintaining credit relationships with a single bank (Mayer, 1994). In such conditions, no bank will confidently incur the fixed costs associated with monitoring without requesting high interest rates.

In countries where share ownership is widely scattered, such as the US and the UK, stock-market listing and recourse to the bond market appear to be the main solution to the funding problems of the corporate sector. This practice does not foster the realization of long-term investment projects for several reasons: (i) the possibility of takeovers tends to limit the time horizon over which both owners and managers of firms plan their activities; (ii) arm's-length sources of financing are at an informational disadvantage when compared with banks, so that the amount of external financing will be lower and its cost possibly higher.

This is consistent with the evidence that smaller firms with well-established relationships with their lenders are less likely to be liquidity-constrained. [1]

While the empirical findings covering various countries seem to agree that close relationships contribute to an increased availability of capital to firms (with the single notable exception of Houston and James (1996)), their effect on the cost of credit is less clear.

According to the "delegated monitoring" argument (Diamond, 1984), delegating an insider bank as the primary monitor and governor of corporate behavior can reduce the costs associated with asymmetric information and moral hazard. Improved knowledge of the borrower's creditworthiness would in turn reduce the riskiness of the loan. Accordingly, the cost of lending should decrease.

[1] See Petersen and Rajan (1994), Weinstein and Yafeh (1994) and Bianco (1997).

On the other hand, both Sharpe (1990) and Rajan (1992) point out that long-term relations provide the bank with such an informational advantage that the strength of competition from other intermediaries or from the financial market is reduced. The bank can choose to use its monopoly power to squeeze the profits of the informationally captive firm by making continuation of its support conditional on higher interest rates, thereby breaking the implicit commitment. According to Rajan (1992), von Thadden (1992) and Weinstein and Yafeh (1994), such a cut in the surplus generated by the project could distort the investment decisions of the "informationally captured" firm, thereby reducing the welfare effect of borrowing from a main bank rather than from multiple sources – either many banks or the capital market.

In principle, the bank could refrain from extracting any rent from its credit relationships in order to discourage customers from seeking alternative, cheaper sources of financing. However, this threat is not credible if the information generated during the relationship cannot easily be acquired by new lenders, as implied by the hold-up hypothesis. [2] As suggested by Sharpe, a more plausible reason not to exploit a monopolistic position could be the bank's concern for its own reputation, which could be tarnished at the eyes of potential customers.

The extent of banks' sensitivity to such arguments remains to be determined empirically.

In some countries businesses have multiple banking relationships and borrow from public markets. Diamond (1991) and Houston and James (1996) suggest that large companies can easily diversify their external financing because they do not require the screening and monitoring activities implicit in a long-term credit relationship. On the contrary, small firms do not have the reputation necessary to obtain credit from different sources and have to rely on a single bank.

Multiple sources of financing can mitigate the hold-up problem. On the other hand, they can affect both the cost of capital and the quality of the investment project in different ways. Each lender will invest fewer resources in monitoring the borrower than it would in the case of a long-term relationship. In these circumstances deserving investment projects might be rejected and firms might use their bank debt to engage in risky activities without being properly monitored by the credit system [3].

[2] Assuming instead that banks learn nothing about borrowers over time and that they behave as credible institutions, Boot and Thakor (1994) present a model where the option for borrowers to switch to another bank is credible, allowing them to obtain a lower interest rate from the current bank.

[3] Using Italian data, Foglia et al. (1998) show that multiple banking relationships weaken the discipline exercised by the banks, thereby increasing the riskiness of the borrowers.

Moreover, in the absence of a long-term commitment, no bank will be induced to act as a risk sharer in periods of financial distress and a "winner's curse" problem could arise. In any case, the restructuring of loans and rescue operations will be more troublesome.

To date empirical evidence concerning the correlation between the degree of competition and the strength of bank–borrower relationships on the one hand, and interest rates on the other hand has provided mixed results.

Taking data collected in 1987 from a sample of 1389 American small firms, Petersen and Rajan (1994) regress the interest rate on the most recent loan on some control variables and two relationship variables: the length of the relationship and the number of banks from which each firm borrowed. They find most of the control variables and the length variable to be not significant, while the number of banks is revealed as having a significantly positive correlation with the interest rate. They interpret this result as showing that exclusive relationships reduce the cost of credit. In a 1995 paper, based on 1277 firms drawn from the same sample, they relate the interest rate with the age of firm, estimating different age intercepts and slopes for different levels of the concentration index of the market for deposits, the latter being used as a proxy for the strength of the lending relationship. They find a negative correlation between the interest rate and the firm's age for each level of the concentration index, but the coefficient has a lower magnitude in more concentrated markets. They interpret this result as evidence of the existence of the hold-up problem.

Using data on 371 firms from the same sample but focusing on loans provided only through lines of credit, Berger and Udell (1995) regress the interest rate on the most recent loan on some control variables and the length of the relationship. They find none of the explanatory variables to be significant, except the length variable with a negative sign, showing how durable bank relationships benefit the borrowers.

Weinstein and Yafeh (1994) find only indirect evidence that Japanese firms pay higher interest rates in exchange for the stability of credit.

We also focus our attention on the cost of bank credit. Unlike the previous papers, however, we use a unique data-set covering the relationships for eight years between 2300 large Italian sized firms and 83 banks; the data report the amount and the interest rate for each loan contract, allowing an accurate description over time of each bank–borrower relationship and the degree of competition in the loan market.

3. Credit relationships in Italy

As noted earlier, the distinction between banking and market economies seems to fit the Italian context only partially. An international comparison shows that in Italy borrowing from credit intermediaries accounted for 42% of the stock of external finance of the production sector in 1996, higher than in France, the UK and the US, but lower than in Germany and Japan (Fig. 1).

By international standards, the contribution of risk capital to the financing of the corporate sector in Italy also occupies an intermediate position. However, the limited role of the security market makes it difficult for businesses to consider equity financing as a close substitute for bank debt. Stock market capitalization is lower than in other leading economies, at 21% of GDP on average in the 1993–1997 period, although a large number of companies satisfy minimum capital and profitability requirements for admission to the stock exchange.

For many years the limited development of the stock exchange was attributed to the weakness of demand for shares, which in turn was due to the small share of financial savings entrusted to institutional investors, such as investment funds, pension funds and insurance companies and to the adverse tax treatment of stock holdings. However, funds in the hands of institutional investors have recently grown rapidly, and a tax reform in 1994 places earnings on investment in shares on a similar footing with securities.

As far as the corporate sector is concerned, according to Pagano et al. (1998), companies do not seek listing mainly because they are unwilling to accept the supposed disadvantages of visibility and accountability for their decisions.

In practice, for medium and large sized Italian firms multiple sources of financing are equivalent to business relationships with a large number of banks.

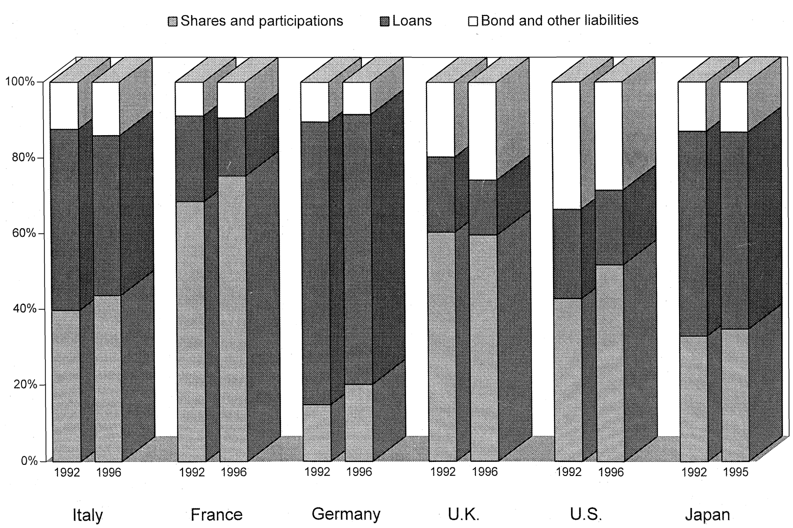


Fig. 1. Financial liabilities of enterprises (percentage composition).

This basic fragmentation of credit relationships in Italy was already described more than 30 years ago. [4] It has attracted renewed attention in connection with the difficulties encountered by some major Italian firms during the 1990–93 recession and important changes in bank legislation aimed at fostering a model of universal banking.

There are several reasons why both firms and banks have found it convenient to perpetuate a system of multiple-credit relationships. As already mentioned, multiple credit relationships can help firms to solve the hold-up problem and to lower the average cost of debt. From the banks' point of view, the system could be seen as a sort of "insurance policy", in which the risks are shared among banks and the involvement of other lenders reduces individual losses if firms fall into financial distress.

Structural and institutional factors may also be responsible for credit fragmentation. The small average size of Italian banks and the need to avoid excessive concentration of risk may have induced banks to share the financing of large borrowers. The operational and maturity specialization imposed on banks between 1936 and 1993 contributed to limit the amount of credit extended by individual banks and the time horizon of their financing, which consisted mainly of overdraft facilities. This could have induced firms to maintain relationships with numerous banks in order to stabilize overall credit flows over time.

Using the data held by the Italian Central Credit Register, [5] Padoa-Schioppa (1993) has shown that, on average, in 1992 fewer than two banks financed firms whose bank debt was less than $658,000 (Lit 1 billion). [6] These firms accounted for 72% of banks' corporate customers and 12% of total bank loans. Firms with bank debt over $20 million (30 billion lire) (less than 1% of corporate borrowers but one third of total bank loans), maintained credit relationships with more than ten banks each. [7]

Given the evidence that the multi-banking phenomenon is stronger for large borrowers, we focus on a sample of major firms to look more deeply at the choice between one main bank and several credit relationships.

In particular, we study the borrower–lender relationship using a subset of around 3000 large and medium-sized firms taken from those recorded in the

[4] Reference to such a situation is made in the Governor's "concluding remarks" contained in Banca d'Italia, Annual Report for 1963.

[5] The Central Credit Register (Centrale dei rischi) is a data-set held by the Bank of Italy since 1962 which records all individual bank loans above a given threshold. A sample of 83 banks, which includes the largest ones, is also required to report interest rates on individual short-term loans.

[6] In the paper we use an exchange rate between Italian Lira and US Dollar equal to 1520.

[7] For the 250 large listed companies considered by Houston and James (1996) the average number of banking relationships is five.

Central Credit Register between 1985 and 1994. The sample comprises the industrial and holding companies with the highest amounts of bank borrowing.

For each credit relationship, the data-set provides information on the type and volume of credit lines and the interest rates on short-term loans. For the only purposes of this work firms have been divided into three groups according to their level of bank debt (see Table 1):

Firms	Bank debt ($ million)
"Major"	>130
"Large"	33–130
"Medium-sized"	13–33

A number of indicators are used to measure the overall relations between individual firms and the banking system: (i) the number of banks from which a firm borrows; (ii) the fraction of borrowing granted by the largest single lender; (iii) the minimum number of banks required to get 50% of total bank debt and (iv) the degree of concentration of borrowing amongst lenders.

Table 2 describes the multiple-bank credit relationships for three years: 1985, 1990 and 1994. With reference to 1994, it can be seen that:

1. major firms borrow from an average of 30 banks compared with 11 for medium-sized businesses; these firms' largest lender account for 29% of their total bank debt compared with 40% for medium-sized borrowers; 50% of major firms' total loans are granted by 4.1 banks, compared with 2.8 for medium-sized firms: the Herfindhal index of their credit relationships is 0.18, whereas it is 0.31 for medium-sized firms;

2. focusing on the firms for which the Herfindahl ratio is higher than the group mean, it is apparent that even in our sample biased towards large firms many borrowers maintain relatively close credit relationships with individual banks. For one third of the companies included in the sample the average number of bank lenders is only slightly higher than one. The presence of these borrowers makes the hold-up problem relevant even for the companies included in our sample;

3. the firms with a low concentration of bank borrowing tend to make relatively little use of individual credit lines;

4. a tendency towards a reduction in the fragmentation of credit relationships is clearly observable. It is connected more with the consolidation process that is developing in the Italian banking system and with the cyclical slowdown in the demand for credit than with a change in the behavior of either the banks or the borrowers.

Our sample can also be used to monitor the stability of lending relationships over time. One way is to compare relationships at the beginning and at the end of the period. For the 1500 firms recorded in the Central Credit Register in each year from 1985 to 1993, 41% of the relationships reported in 1993

Table 1
Sample size and firms' debt

Firm size	1985				1990				1994			
	No. of firms	Credit granted ($ mls)	Credit drawn ($ mls)	No. of banks	No. of firms	Credit granted ($ mls)	Credit drawn ($ mls)	No. of banks	No. of firms	Credit granted ($ mls)	Credit drawn ($ mls)	No. of banks
Major	71	724	470	43	155	757	482	40	177	878	588	30
Large	180	121	61	25	471	112	60	21	572	95	61	16
Medium-sized	1911	16	7	12	1855	28	13	13	1473	28	16	11
All	2162	48	27	14	2481	89	51	16	2222	113	73	14

Source: Italian Credit Register and Balance Sheet Register.

Table 2
Charateristics of credit relationships by firms' size and degree of loan concentration

	Major firms			Large firms			Medium-sized firms		
	All	High concentration	Low concentration	All	High concentration	Low concentration	All	High concentration	Low concentration
1985									
No. of firms	71	20	51	180	53	127	1911	602	1309
Bank loans ($ mls.)									
Granted (*a*)	724	516	805	121	101	129	16	8	20
Drawn (*b*)	470	386	504	61	64	60	7	5	9
Share of credit drawn (*b*/*a*%)	64.9	74.8	62.6	50.4	63.4	46.5	43.8	62.5	45.0
No. of banks per firm	43	29	49	25	14	30	12	5	15
Share % of the first bank	23.2	43.0	15.4	24.2	44.2	15.9	36.3	66.4	22.4
No. of banks summing up to the first 50% of total loan granted	5.2	2.3	6.3	5.0	1.9	6.3	3.3	1.3	4.2
Herfindahl index of concentration	0.123	0.258	0.070	0.139	0.291	0.076	0.267	0.568	0.129
1990									
No. of firms	155	47	108	471	143	328	1855	573	1282
Bank loans ($ mls.)									
Granted (*a*)	757	567	839	112	89	123	28	17	32
Drawn (*b*)	482	412	513	60	58	61	13	9	14
Share of credit drawn (*b*/*a*%)	63.7	72.7	61.1	53.6	65.2	49.6	46.4	52.9	43.8
No. of banks per firm	40	25	47	21	12	26	13	5	16
Share % of the first bank	24.4	44.6	15.6	27.5	50.9	17.3	36.8	69.5	22.1
No. of banks summing up to the first 50% of total loan granted	5.1	2.1	6.4	4.5	1.6	5.8	3.4	1.2	4.3

Herfindahl index of concentration	0.133	0.276	0.071	0.172	0.365	0.088	0.273	0.602	0.126
1994									
No. of firms	177	48	129	572	177	395	1473	474	999
Bank loans ($ mls.)									
Granted (*a*)	878	449	104	95	66	108	28	16	34
Drawn (*b*)	588	349	68	61	60	62	16	13	17
Share of credit drawn (*b*/*a*%)	67.0	77.7	65.4	64.2	90.9	57.4	57.1	81.3	50.0
No. of banks per firm	30	17	35	16	9	20	11	6	14
Share % of the first bank	29.4	56.4	19.3	36	66.5	22.4	40.2	71	25.5
No. of banks summing up to the first 50% of total loan granted	4.1	1.6	5.1	3.4	1.3	4.3	2.8	1.2	3.6
Herfindahl index of concentration	0.184	0.422	0.096	0.265	0.584	0.123	0.309	0.641	0.152

Source: Italian Credit Register and Balance Sheet Register.

(accounting for 26% of total borrowing) were established with different banks from those recorded in 1985. In 1993, 32% of the credit relationships recorded in 1985, totaling 20% of bank debt at that time, were no longer recorded. During this period, the number and size of new banking relationships were larger than those of interrupted relationships. This suggests that firms attempted to broaden the range of financial sources rather than substitute one bank with another. [8]

To sum up, bank relationships are more fragmented in Italy than in other leading economies (Conigliani et al., 1997). However, for a certain range of firms multiple-credit relationships do not preclude the existence of larger and relatively stable ties with an individual bank. In the case of medium-sized firms, in particular, the level of concentration is so high that the possibility that some of them have been captured by a bank must be considered.

4. An econometric study of factors affecting interest rates on bank loans

4.1. The sample

In order to ascertain empirically the effect of credit relationships on the cost of borrowing, an econometric analysis was performed, matching the data on the credit portfolios of individual banks with data on the firms' balance sheets. The balance sheet data are registered in the Balance Sheet Register (Centrale dei bilanci), held by a company jointly owned by the Central Bank and Italian major banks, in which participating banks record the balance sheets of their corporate customers.

Firms whose balance sheet and interest rate data were incomplete and inconsistent were dropped from the original sample of 3000 firms. This selection process leaves us with an unbalanced panel of 2331 firms, with an average of around 1400 per year; 10% of the 2331 firms are classified as insolvent borrowers at some point during the sample period. The data-set includes for each firm information on their balance sheets, borrowing vis-à-vis individual banks and the interest rates charged by 83 banks representing two thirds of the Italian banking system.

As already mentioned, the sample is biased towards large firms: during the period 1985–1993, on average 57% of the industrial firms had more than 200 employees, only 13% had less than 50 employees. Among holding companies, 40% had total assets of more than $200 million. Table 3 reports summary

[8] Another more direct way to control the stability of lending is to consider the number of years each relationship lasts. This is the procedure we follow in the regression analysis (see discussion below).

Table 3
Bank and firm summary statistics ($ million)

	1987						1994					
	No. of obs.	Min	Median	Max	Mean	Std	No. of obs.	Min	Median	Max	Mean	Std
Banks												
Total assets (BSIZE)	83	136	2.532	40.089	5.943	8.361	75	165	4.736	96.319	13.034	19.790
Loans	83	54	827	11.432	1.786	2.343	75	65	1.565	28.191	4.409	4.409
Deposits	83	114	1.506	19.217	3.251	4.104	75	133	2.028	40.260	5.950	8.489
Operating costs	83	6	71	1.323	176	250	75	7	119	2.303	362	362
Bad and non-performing loans	83	4	46	897	140	198	75	5	84	3.971	389	389
	1986						1993					
Industrial firms												
Total revenues (FSIZE)	1.318	0.001	34	11.918	118	551	1.051	0	71	20.220	212	927
No. of employees[a]	1.316	0	223	114.990	809	4.619	1.000	0	277	105.835	928	5.386
Holding companies												
Total assets (FSIZE)	45	5	227	13.798	859	2.259	64	29	147	24.238	1.166	3.445
No. of employees [a]	45	0	25	1.239	110	227	59	0	21	3.409	193	609

[a] Some firms did not report the number of employees.

statistics for banks and firms at the beginning and the end of the period. In 1993, the average industrial firm employed more than 900 workers and half of them more than 280 employees. The average holding company had total assets of $1 billion (1700 billion lire); for 50% of them the value of assets was more than of $150 million. Most of these are privately owned and located in the northern regions of Italy.

The size of the banks is also large: in 1994 the average bank had assets of $13 billion; their overall short-term loan market share was around 75%.

4.2. The specification

Our empirical analysis differs from previous empirical research for a number of reasons: (i) our data are not drawn from a survey targeted on firms but are regularly supplied by Italian banks to the Credit Register held by the Central Bank and to the Balance Sheet Register; (ii) the data cover several years, allowing a more complete econometric analysis; (iii) our data allow an accurate description of individual lending relationships and of the observable credit risk of the borrower.

The dependent variable is the spread between the interest rate charged to the individual borrower by each bank in the last quarter of the year and the average interest rate on Treasury bills at 3, 6 and 12 months at issue (SPREAD). The loan interest rate is referred to short-term lending, mainly granted in the form of uncollateralized overdraft facilities.

The explanatory variables are divided into three basic categories:

1. variables pertaining to the banks' operating characteristics;
2. variables relating to borrowers' economic and financial situations;
3. variables describing the credit relationships of each firm.

To avoid endogeneity problems, the firm-specific variables refer to one year lagged values, and bank variables are computed as yearly average data.

As for the banks, the following variables have been considered:

- total assets (BSIZE): large banks tend to have stronger bargaining power and to charge high interest rates (positive expected sign);
- the share of bank deposits invested in loans (LDEP): a positive correlation is expected since banks with higher loan-to-deposit ratios got them by expanding the set of firms to whom they lend. This implies that on the margin, they have lent to higher risk borrowers [9];

[9] Another reason to expect a positive sign is that banks with a lower loan-to-deposit ratio usually aim at shifting the composition of their assets towards the most profitable item (loans), and are prepared to charge low interest rates. This is what happened in Italy after the removal of the credit ceiling, which was in force from 1973 to 1983 for monetary policy purposes (Cottarelli et al., 1986).

- the ratio between bad and non-performing debts and total bank loans (BADL): banks with a high ratio tend to charge high interest rates in order to replenish their provisions (positive expected sign);
- operating expenses as a ratio of total assets (COST): the banks with high production costs tend to charge high interest rates to achieve satisfactory net earnings (positive expected sign);
- the ratio between bank's capital and total loans (SOLVENCY). The expected effect of this variable is negative: highly capitalized banks tend to lend to less risky firms since owners have more at stake;
- credit losses two years ahead (LOSSES2), to control for banks' risk preferences (as in Berlin and Mester, 1998): a positive sign is expected since banks that are now lending to riskier borrowers at higher interest rates will show greater losses in the future.

The second group of explanatory variables includes some indicators of the profitability and soundness of corporate borrowers:

- return on assets (ROA), positively related to SPREAD, signaling high risk borrowers with an average higher return requiring a higher interest rate;
- a liquidity ratio (short-term assets over short-term liabilities, LIQ): a high ratio denotes a strong ability to face sudden liquidity needs and should be reflected in lower interest rates paid to lenders (negative expected sign);
- an equity to debt ratio (the volume of own funds vis-à-vis the firm's indebtedness to others, LEV): a high value of this ratio enhances creditworthiness (negative expected sign);
- a proxy variable for the debt service capacity, defined as the ratio between interest expenses and the gross operating margin (INTCOST): a high value of this ratio could be interpreted as indicating a borrower's financial fragility and make lending riskier (positive expected sign);
- the firm's size (FSIZE), proxied by total revenue for industrial companies and by total assets for holding companies: the borrower's bargaining power is directly correlated with its size (negative expected sign);
- a proxy for the relative size of both lender and borrower, defined as the ratio of BSIZE over FSIZE (RELSIZE) (positive expected sign).

The third group of independent variables describes the relationship between each borrower and each lender. The strength of a lending relationship can be measured by two factors: the extent to which a lender is involved in the financing of a borrower and the stability of its financial support. The following variables have been considered:

- the volume of the credit facility granted to each customer (GRA), a size index of the credit relationship;
- the share of each bank over the total credit lines granted by the banking system to individual borrowers (BSHARE), as a measure of a bank's importance for a firm;

Table 4
Interest rate equation (Variable summary statistics)

Variables	1987						1994					
	No. of obs.	Min	Median	Max	Mean	Std	No. of obs.	Min	Median	Max	Mean	Std
Dependent variable												
SPREAD (%) (spread over Treasury bonds)	14.096	−1.87	1.06	11.34	1.37	1.78	15.912	−3.40	−0.16	12.04	0.76	2.44
Bank variables												
LDEP = loans/deposits	83	0.313	0.544	0.885	0.535	0.106	75	0.404	0.719	1.035	0.714	0.713
COST = Operating expenses/total assets	83	0.019	0.031	0.052	0.031	0.007	75	0.016	0.029	0.056	0.030	0.007
BADL = bad and non performing loans/total loans	83	0.016	0.066	0.351	0.077	0.049	75	0.012	0.057	1.002	0.087	0.087
SOLVENCY = capital/total loans	83	0.062	0.195	0.529	0.209	0.098	75	0.090	0.200	0.438	0.215	0.072
LOSSES2 = losses on loans two years later	83	0.299	8.721	336.498	27.010	48.526	68	0.925	29.213	853.167	91.363	148.98
	1986						1993					
Industrial firm variables												
ROA (return on assets)	1318	−0.147	0.096	0.389	0.104	0.072	1.051	−0.144	0.066	0.392	0.067	0.067
LIQ = current assets/current liabilities	1318	0.243	1.233	3.612	1.320	0.428	1.051	0.208	1.093	3.768	1.157	0.415
INTCOST = interest expenses/gross operating margin	1318	0.005	0.425	2.000	0.580	0.503	1.051	0.000	0.641	2.000	0.809	0.572
LEV = equity/debt	1318	−0.066	0.613	4.986	0.924	0.901	1.051	−0.099	0.436	4.960	0.700	0.804

Holding company variables												
ROA (return on assets)	45	−0.123	0.084	0.284	0.088	0.078	64	−0.149	0.054	0.215	0.042	0.072
LIQ = current assets/current liabilities	45	0.000	0.841	4.519	1.125	1.173	64	0.000	0.527	3.905	0.692	0.699
INTCOST = interest expenses/financial revenues less labour costs	45	0.154	0.743	2.000	0.851	0.570	64	0.223	1.262	2.000	1.341	0.572
LEV = equity/debt	45	0.027	0.856	4.529	1.304	1.198	64	0.013	0.808	4.317	0.999	0.824
	1987						1994					
Relationship and concentration variables												
NQUART (lenght of relationship in quarters)	14.096	1.0	10.0	10.0	9.1	2.1	15.912	1.0	32.0	38.0	28.0	11.4
GRA (credit granted by each bank to each customer in \$ millions)	14.096	0.030	1.316	411	3.668	11	15.912	1.32×10^{-3}	2.450	1175.16	8.551	34.697
Table 4 (Continued) BSHARE = bank loan/firm's total bank loans	14.096	6.00×10^{-5}	0.039	1.000	0.060	0.065	15.912	9.9×10^{-6}	0.041	1.000	0.066	0.088
FSHARE = bank loan/bank's total loans	14.096	1.6×10^{-6}	0.0002	0.048	0.001	0.002	15.912	6.5×10^{-8}	0.0001	0.085	0.001	0.002
DRAWNSH = credit drawn /credit granted	14.096	1.00×10^{-5}	0.349	9.560	0.403	0.353	15.912	2.00×10^{-5}	0.580	390.600	0.640	3.240
NBANKS (no. of banks per firm)												
Industrial firms	1.318	1.0	14.0	159.0	17.0	13.0	1.051	2.0	15.0	159.0	16.0	10.0
Holding companies	45	1.0	20.0	131.0	27.0	26.0	64	2.0	16.0	78.0	19.0	15.0

- loans to individual borrowers as a ratio of the bank's total loan portfolio (FSHARE), a means of measuring the importance of the customer for the bank;
- the share of the credit facility actually drawn down at each bank (DRAWNSH), as an indicator of the tightness of the credit relationships. Table 2 and Foglia (1995) show that this ratio is higher for firms with a higher concentration index, a larger share of credit borrowed at the main bank and a lower number of lending banks.

The length of the relationship is measured by the number of quarters each firm has been dealing with each lender (NQUART) [10].

If the hold-up hypothesis holds, the interest rate would be positively linked to the relationship variables. If instead the bank is not willing to extract any rent, the sign of these variables would be negative.

Multiple sources of financing should mitigate the hold-up problem, while highly concentrated lending relationships should mitigate the threat of competition, leading to higher individual interest rates. To take into account the influence of competition on the pricing policy of each bank, the number of banks lending to the same borrower was introduced in the specification (NBANKS, negative expected sign) [11].

A concentration ratio of local markets for bank loans has been used to control for the different intensity of bank competition in the four major areas (north-east, north-west, center, south and islands) of Italy (BHF) (positive expected sign) [12].

The final model is as follows:

$$\mathrm{SPREAD}_{ij} = f(\mathrm{BV}_i, \mathrm{FV}_j, \mathrm{RV}_{ij}, \mathrm{NBANKS}_j, \mathrm{BHF}), \tag{1}$$

where $i = 1, \ldots, n$ ($n =$ number of banks $= 83$), $j = 1, \ldots, m$ ($m =$ number of firms $= 2331$), BV is bank variables, FV firm variables, RV relationship variables.

The model was tested over a seven-year span, from 1987 to 1994 (see note 10), divided equally between a phase of economic expansion and a period of serious recession. The number of observations is around 120,000. Tables 3 and

[10] We began tracking the length of each relationship from September 1985, the start of our sample period; this has the obvious drawback of considering all relationships as beginning at the same time. To allow the variable NQUART some cross-section variability, we performed regression analysis only from 1987 on.

[11] Petersen and Rajan (1994) include only the length variable and interpret the number of banks per firm also as an indirect measure of the closeness of a relationship. Berger and Udell (1995) include only the length variable.

[12] This variable is the Herfindahl index of the market shares of bank loans to firms for each geographic area.

4 show the distribution of observed values of all variables included in the model.

5. Empirical results

The interest rate equation has been estimated on both cross-section data and the pooled data-set using panel techniques.

5.1. Cross-section estimates

In addition to the variables listed above, several other dummies were used in the cross-section regressions:

- dummy variables (DAREA) for the four geographical areas, to control for the higher degree of credit risk usually associated with southern firms (positive expected sign for the South; negative expected sign for other areas);
- a dummy variable for firms belonging to large industrial groups with strong bargaining power (DGROUP) (negative expected sign);
- a dummy variable to distinguish between private and State-owned enterprises (DPRIV); for a long time the latter were considered less risky, in the mistaken belief that there exist an implicit unlimited government guarantee (positive expected sign);
- a dummy variable to distinguish between industrial and financial holding firms (DFIN): the latter being the parent firms of large industrial groups, with a strong bargaining power (negative expected sign);
- dummy variables for the 23 industrial classifications of firms' economic activity, generally characterized by different degrees of risk (DBRANCH).

The results of separate cross-section estimates for each year from 1987 to 1994 are shown in Table 5. Most of the explanatory variables behave consistently with the hypotheses and significantly affect the interest rates charged to borrowers.

Among the variables that describe banks, COST and BADL are significantly related to the dependent variable over the whole time span.

As regards the firms variables, LEV and INTCOST are always significant.

All the variables describing credit relationships, except NQUART, significantly affect the dependent variable in each cross-section period. The NBANKS coefficient shows that doing business with a large number of banks allows a borrower to pay lower interest rates to each. On the other hand, according to the GRA, BSHARE and FSHARE coefficients, the higher the share of total debt financed by a single bank, the lower the interest rate charged by that bank to the firm. The NQUART coefficient is significant only twice; in 1992, when its cross-section variability is more plausible, it has a negative sign. We discuss these findings in detail below.

Table 5
Cross-section estimates (one year lagged firms' balance sheet variables)

Explanatory variables	Number of periods in which the coefficient is significant at 5% level	Sign
Bank variables		
LDEP	7	+ negative in 1990
COST	8	+
BADL	8	+
BSIZE	7	+
SOLVENCY	5	– positive in 1991 and 1994
LOSSES2	6	+ negative in 1988 and 1991
Firm variables		
$FSIZE_{(-1)}$	6	+
$ROA_{(-1)}$	3	Negative in 1987–1988 and positive in 1994
$LIQ_{(-1)}$	4	–
$INTCOST_{(-1)}$	8	+
$LEV_{(-1)}$	8	–
$RELSIZE_{(-1)}$	2	Positive in 1989 and negative in 1992
Relationship variables		
NQUART	2	Positive in 1987 and negative in 1992
GRA	8	–
BSHARE	8	–
FSHARE	8	–
DRAWNSH	7	Negative in 1987–1989 and positive in 1991–1994
Concentration variables		
NBANKS	8	–
Dummy variables		
DGROUP	6	–
DAREA	8	Negative for North, positive for South
DFIN	4	+ negative for 1987
DPRIV	6	Positive in 1987–1989, negative in 1991–1993

Dependent variable: SPREAD; time span: eight periods, from 1987 to 1994.

5.2. Pooled CS/TS estimates

Panel estimates make it possible to exploit the data more efficiently and to take into account factors affecting individual banks and borrowers which are not modeled explicitly through the variables included in the regression equation. Given the large size of the sample, we estimated a fixed-effects model, that models individual heterogeneity through variables whose values differ among the agents but do not change over time (individual-time invariant variables).

We also included variables that stay constant across cross-sectional units but vary through time (period-individual invariant variables). Considering only significant variables, the estimated regression is as follows:

$$\begin{aligned}\text{SPREAD} = {} & \alpha_1\text{LDEP} + \alpha_2\text{COST} + \alpha_3\text{BADL} + \alpha_4\text{BSIZE} \\ & + \alpha_5\text{SOLVENCY} + \alpha_6\text{LOSSES2} + \alpha_7\text{FSIZE}_{(-1)} \\ & + \alpha_8\text{ROA}_{(-1)} + \alpha_9\text{LIQ}_{(-1)} + \alpha_{10}\text{INTCOST}_{(-1)} + \alpha_{11}\text{LEV}_{(-1)} \\ & + \alpha_{12}\text{BHF} + \alpha_{13}\text{NQUART} + \alpha_{14}\text{GRA} + \alpha_{15}\text{BSHARE} \\ & + \alpha_{16}\text{FSHARE} + \alpha_{17}\text{NBANKS} + \alpha_{18}\text{DRAWNSH} + \varepsilon. \end{aligned} \tag{2}$$

To account for fixed individual and temporal effects, we added a dummy variable for each firm, bank and time period considered. The dummies relating to the firms give an important contribution in terms of reducing the error sum of squares of the model, whereas this is not the case for those referring to individual banks. This may indicate that differences between banks are well captured by the bank variables.

The results of the regression including only firms and time fixed effects are shown in Table 6. Among the three groups of explanatory variables, those describing the bank's operating characteristics and the structure of the firm's loans are the most significant.

The bank's size (BSIZE) is positively correlated with the interest rate on loans, but its influence is very limited: a bank with assets of about \$20 billion (one standard deviation away from the 1994 sample mean) more than another can impose an interest rate that is only 0.005 basis points higher. The LDEP variable turns out to be quite important: a bank whose ratio is near to the average value (70% in 1994) will charge a borrower an interest rate 6.5 basis points higher than a bank with LDEP around 60%. The ratio of bad debt to total credit, BADL, is positively related to the interest rate on loans and its effect is also large: for example, a bank with a bad debt ratio one standard deviation higher than the sample mean (9% in 1994) will charge customers an interest rate 22 basis points higher. Finally, a ratio of operating expenses to total assets, COST, of 0.007 higher than the sample mean causes an interest rate 18 basis points higher. The ratio between capital and total loans (SOLVENCY) is negatively related to the dependent variable: a bank with a ratio 7 percentage points (one 1994 standard deviation) higher than the sample mean will charge interest rates 4 basis points lower. The sign of the LOSSES2 is positive but the impact is negligible.

The sets of variables relating to borrowers provide a smaller contribution to the estimate. The size of the borrower, FSIZE, is negatively related to the interest rate but its effect is extremely small. The LEV variable shows a negative relationship with the SPREAD: a firm with an equity to debt ratio twice the average will be charged an interest rate two basis points lower. The liquidity

Table 6
Fixed effects panel estimates (one year lagged firms' balance sheet variables; firm and time effects)

Explanatory variables	Coefficient	t statistic	Pr > ITI	Partial sum of squares	%
Bank variables					
LDEP	0.655	13.50	0.0001		
COST	25.267	26.04	0.0001		
BADL	2.425	22.64	0.0001		
BSIZE	2.62×10^{-6}	11.25	0.0001		
SOLVENCY	−0.596	−6.54	0.0001		
LOSSES2	1.80×10^{-7}	10.32	0.0001	4.299	2.01
Firm variables					
$\text{FSIZE}_{(-1)}$	-5.02×10^{-5}	−4.16	0.0001		
$\text{ROA}_{(-1)}$	0.407	2.69	0.0072		
$\text{LIQ}_{(-1)}$	−0.081	−4.10	0.0001		
$\text{INTCOST}_{(-1)}$	0.226	13.21	0.0001		
$\text{LEV}_{(-1)}$	−0.0297	−2.49	0.0127	579	0.27
BHF	10.688	6.23	0.0001	101	0.05
Relationship and concentration variables					
NQUART	0.0062	7.27	0.0001		
GRA	−0.0015	−5.94	0.0001		
BSHARE	−3.302	−34.00	0.0001		
FSHARE	−44.772	−9.32	0.0001		
DRAWNSH	−0.101	−14.21	0.0001		
NBANKS	−0.0135	−11.76	0.0001	4.358	2.04
Fixed effects		f value	PR > F		
Firm		24.85	0.0001	143.923	67.32
Time		3728.99	0.0001	60.541	28.32
Model		40.53	0.0001	213.801	100.00
R^2	0.447				

Dependent variable: SPREAD; time span: eight periods, from 1987 to 1994.

ratio (LIQ) significantly affects the interest rate charged on bank loans: for an industrial firm, a one standard deviation reduction in the liquidity ratio is associated with an increase in the interest rate of 3.3 basis points. The debt service capacity, proxied by INTCOST, has the predicted sign: an improvement in the ratio from 80% to 20% allows a 13 basis points reduction in the interest rate. The ROA variable presents a positive sign, which seems to confirm that banks charge higher interest rates to riskier borrowers.

The variables capturing the strength of the firm–creditor relationships and the degree of competition are highly significant. The evidence shows that closer firm–bank ties reduce borrowers' lending cost.

The share of one firm's debt supplied by an individual bank, BSHARE, is the most significant explanatory variable. The negative sign confirms that it is worthwhile for the firm to strengthen its lending relationship with a bank. A 1% increase in the bank's share gives rise to a reduction in the spread of 3.3 basis points. The actual distribution of the variable (Table 4) ranges from very low values to 100%. The average fraction in 1994 was 7%; for half of the relationships, the fraction of all bank borrowings coming from each bank is below 4%. A firm that increases the fraction of lending by a bank from 7% (the sample mean bank share) to 35% (the largest bank's average share, see Table 2) can expect to pay almost one percentage point less.

On the other hand, increasing the length of the relationship, NQUART, does not seem to favor the borrower, consistently with the empirical results presented by Petersen and Rajan (1994). However, the impact on interest rates is small: a further one-year extension of a relationship with the same bank brings about an increase in the interest rate of 2.5 basis points [13]. The weak result for this variable might be connected with the difficulty of proxying the actual length of the credit relationships on the basis of the time span for which data are available. The length of the NQUART variable is left and right censored and the only loans for which the length of the relationship is known for sure are those which began after 1985 and ended before 1994: only 100 out of more than 120.000 relationships in the sample. By limiting the estimation procedure to these relationships, both the positive sign of the variable and its negligible impact are confirmed [14].

The impact of the amount of financing granted by each bank to a customer, GRA, is minimal. Rather, lending to an individual firm as a percentage of the bank's total credit portfolio (FSHARE) is negatively related to the interest rate, which decreases by 9 basis points if that share goes up by one standard deviation. As expected for the DRAWNSH variable, leaving a substantial share of unused facilities leads to a higher interest rate: a firm using 60% would be charged an interest rate 4 basis points higher than a borrower exploiting the whole credit line.

[13] In other words, it takes 16 months to offset the interest rate reduction due to a 1% increase in a bank's share. To neutralize the beneficial effect of a concentration of lending with a main bank (e.g. by increasing the fraction of loans borrowed from the same bank from 7% to 35%, as in the above example), the relationship with that bank would have to last 40 more years!

[14] In an attempt to improve the results concerning the length of the credit relationships, the whole sample has been divided into four portions: the relationships which began after 1985 and ended before 1994; those that began before 1985 and ended before 1994; those that began after 1985 and ended after 1994; and those that began before 1985 and ended after 1994. Dummy variables for loans in the last three categories were introduced in the cross-section estimates, but their coefficients proved highly unstable over time.

The concentration index of the local market area, BHF, has a positive sign: in less competitive local markets, firms are charged higher interest rates.

The NBANKS variable confirms the beneficial effect of competition on the price of each lending relationship. For a new bank granting credit, the interest rate charged by each bank decreases by 1.3 basis points [15].

In order to complement the assessment of competition, an attempt was made to introduce a concentration index of lending to individual firms, defined as the sum of squares of the banks' shares (HFI). Its introduction brings about a negligible change of the other coefficients, whose signs remain unaltered. HFI is positively related to SPREAD, with significance at 1%; the value of its coefficient is 2.26. For given values of BSHARE and NBANKS, a decrease of HFI implies that lending by other banks is less concentrated. The positive sign means that an increasing competition produces a reduction in the interest rate charged by the bank.

Analyzing the joint effects of HFI, NBANKS and BSHARE gives some indication of the hold-up problem [16]. An increase in the share of each bank over total credit to an individual borrower (BSHARE) has a negative effect on the interest rate if BSHARE is lower than 80%. Higher values of this share would be accompanied by an increase in the interest rate.

6. Conclusions

This paper aims to add to the growing literature on the impact of credit relationships on loan terms offered to borrowers. The reference to Italy is particularly useful, since multiple banking relationships are very prevalent and a large amount of statistical information is available.

The data refer to a sample of individual credit lines granted by 83 Italian banks to 2300 large and medium-sized Italian firms in the period 1987–1994. The database contains the amount and the interest rate for each loan contract,

[15] One borrower can obtain the same interest rate reduction from a bank either by increasing by 1% the fraction of borrowing from that bank or by applying to 2.4 new banks for the same amount of loans. To match the decrease in the rate when concentrating borrowing with a main bank, a firm would need 71 new banks.

[16] To analyze the impact on SPREAD of a variation of BSHARE and/or NBANKS it is necessary to take into account the contemporaneous change of the Herfindahl index of concentration. Using the estimated coefficients of NBANKS and HFI, the derivative of SPREAD with respect to NBANKS is always negative. The derivative of SPREAD with respect to BSHARE is a function of the value of BSHARE itself and it is negative up to a share of 80%. This methodological point is discussed at length in D'Auria and Foglia (1997).

allowing an accurate description of each bank–borrower relationship over time and of the degree of competition in the loan market.

We selected three groups of factors that influence the pricing policies of banks: the operating characteristics of the banks and their risk preferences, the size and riskiness of the borrowers, the structure of credit relationships. A fourth factor is represented by bank competition in local markets.

The results of the panel regression show the closeness of lending relationships measured by a bank's share of customer debt as the main determinant of individual interest rates. Larger shares are associated with lower interest rates. A hold-up problem seems to arise only when the relationship with the main bank is almost exclusive.

The magnitude of the effect is also important: a firm that decides to concentrate borrowing at a main bank can expect to pay an interest rate one percentage point lower than a firm with more fragmented relationships. The influence of the length of the individual credit relationship, though statistically significant, seems very low.

The degree of competition in the credit market also significantly affects interest rates, but once again the impact is very small. Other things being equal, a borrower can expect to pay each bank a slightly lower interest rate if it increases the number of lending banks. We interpret this result as showing that some degree of competition does not weaken the relationship with a main bank; it is used by the borrower as an incentive to discourage the relationship bank from extracting a monopoly rent.

Concentrating the debt with a main bank while having relationships also with other institutions seems like an optimal arrangement for the borrower, since the main bank would have an informational advantage that would allow it to offer lower interest rates, and the multiple lenders would alleviate the potential hold-up problem.

Acknowledgements

We are grateful to Anil Kashyap, George Kaufman, Randall Kroszner, Raghuram Rajan, and two anonymous referees for helpful suggestions on a previous draft. We also thank Elijah Brewer, Douglas Evanoff, Curt Hunter and all seminar participants at the Federal Reserve Bank of Chicago. The views expressed in the article are those of the authors and do not involve the Bank of Italy.

References

Aoki, M. and Patrick, H. (Eds.), 1994. The Japanese Main Bank System, Oxford University Press, Oxford.

Berger, A.N., Udell, G.F., 1994. Did risk-based capital allocate bank credit and cause a credit crunch in the United States? Journal of Money, Credit and Banking 26, 585–628.

Berger, A.N., Udell, G.F., 1995. Relationship lending and lines of credit in small firm finance. Journal of Business 68 (3), 351–381.

Berlin, M., Mester, L.J., 1998. Deposits and relationship lending. Working Paper, No. 18/R. Federal Reserve Bank of Philadelphia, Philaedelphia, PA.

Boot, A.W.A., Thakor, A.V., 1994. Moral hazard and secured lending in an infinitely repeated credit market game. International Economic Review 35 (4), 899–920.

Bianco, M., 1997. Vincoli finanziari e scelte reali delle imprese italiane: gli effetti di una relazione stabile con una banca. In: Angeloni, I., Conti, V., Passacantando, F. (Eds.), Le banche e il finanziamento delle imprese. Il Mulino, Bologna.

Calomiris, C.W., 1993. Corporate-finance benefits from universal banking: Germany and the United States, 1870–1914. Working Paper No. 4408. NBER, Cambridge, MA.

Conigliani, C., Ferri, G., Generale, A., 1997. La relazione banca-impresa e la trasmissione degli impulsi della politica monetaria. Moneta e Credito, giugno 50 (198), 139–165.

Cottarelli, C., Galli G., Marullo-Reedtz, P., Pittaluga, G., 1986. Monetary policy through ceilings on bank lending. Economic Policy 1 (3), 673–694.

D'Auria, C., Foglia, A., 1997. Le determinanti del tasso di interesse sui crediti alle imprese. In: Angeloni, I., Conti, V., Passacantando, F. (Eds.), Le banche e il finanziamento delle imprese. Il Mulino, Bologna.

Diamond, D.W., 1984. Financial intermediation and delegated monitoring. Review of Economic Studies 51, 393–414.

Diamond, D.W., 1991. Monitoring and reputation: The choice between bank loans and directly placed debt. Journal of Political Economy 4, 689–721.

Foglia, A., 1995. Il finanziamento delle imprese in Italia: considerazioni teoriche e aspetti empirici. Banca, Impresa e Società, No. 1.

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

Hoshi, T., Kashyap, A., Scharfstein, D., 1990. The role of banks in reducing the costs of financial distress in Japan. Journal of Financial Economics 27, 67–88.

Houston, J., James, C., 1996. Bank information monopolies and the mix of private and public debt claims. Journal of Finance 51 (1), 1863–1889.

Mayer, C., 1988. New issues in corporate finance. European Economic Review 32, 1167–1189.

Mayer, C., 1994. The assessment: Money and banking: theory and evidence. Oxford Review of Economic Policy 10 (4).

Padoa-Schioppa, T., 1993. Profili di diversità nel sistema bancario italiano. In: Bollettino Economico, Banca d'Italia, No. 22.

Pagano, M., Panetta, F., Zingales, L., 1998. Why do companies go public? An empirical analysis. Journal of Finance 53 (1), 27–64.

Petersen, M., Rajan, R.G., 1994. The benefits of lending relationships: Evidence from small business data. Journal of Finance 49 (1), 3–37.

Petersen, M., Rajan, R.G., 1995. The effect of credit market competition on lending relationships. Quaterly Journal of Economics 110 (2), 407–443.

Rajan, R.G., 1992. Insiders and outsiders: The choice between informed and arm's-length debt. Journal of Finance 47 (4), 1367–1400.

Sharpe, S., 1990. Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships. Journal of Finance 45 (5), 1069–1087.

Sheard, P., 1989. The main bank system and corporate monitoring and control in Japan. Journal of Economic Behavior and Organization 11, 399–422.

von Thadden, E.L., 1992. The commitment of finance, duplicated monitoring and the investment horizon. Working Paper No. 27. CEPR, London.

Weinstein, D.E., Yafeh, Y., 1994. On the cost of a bank centered financial system: Evidence from the changing main bank relations in Japan. Discussion Paper No. 1685, Harvard Institute of Economic Research, Harvard University, Cambridge, MA.

THE JOURNAL OF FINANCE • VOL. LV, NO. 3 • JUNE 2000

Multiple versus Single Banking Relationships: Theory and Evidence

ENRICA DETRAGIACHE, PAOLO GARELLA, and LUIGI GUISO*

ABSTRACT

A theory of the optimal number of banking relationships is developed and tested using matched bank-firm data. According to the theory, relationship banks may be unable to continue funding profitable projects owing to internal problems and a firm may thus have to refinance from nonrelationship banks. The latter, however, face an adverse selection problem, as they do not know the quality of the project, and may refuse to lend. In these circumstances, multiple banking can reduce the probability of an early liquidation of the project. The empirical evidence supports the predictions of the model.

STUDIES OF THE CORPORATE FINANCIAL STRUCTURE began by focusing on the choice between debt and equity (Harris and Raviv (1991)), and only recently have addressed the choice among alternative forms of debt, such as bank debt, public debt, privately placed debt, and others.[1] From this work a theory of bank-firm relationships has emerged, centered on the idea that banks acquire information about the quality of the firm that is not shared by other financial intermediaries.[2] Much of this theory proceeds on the assumption

* International Monetary Fund, Universita' di Bologna, and Universita' di Sassari and Ente Luigi Einaudi, respectively. We wish to thank Marco Da Rin, Mathias Dewatripont, Giovanni Ferri, Andrea Generale, Andrea Gerali, Giorgio Gobbi, Raghuram Rajan, Andrei Shleifer, Alessandro Sembenelli, and participants in seminars at the Chicago FED, Igier, the University of Brescia, the BIS, and in the 1997 Alicante ASSET meeting; the 1st European Corporate Governance Network Conference; the 1997 CEPR Conference on Information and Learning, Financial Intermediation, and the Macroeconomy; the Center for Financial Studies Conference on Credit Risk Management and Relationship Banking; and the 1998 meetings of the American Finance Association for helpful comments on a previous version of this paper. Valuable suggestions were also received from the editor and from an anonymous referee. This paper is part of a project on "Legal Enforcement and the Structure of Bank-Firm Relations." Financial support was provided by MURST. The Federal Reserve Board has kindly provided us with the public release of the 1988 National Survey of Small Business Finances. The views expressed in the paper are the authors' own, and do not necessarily reflect those of the International Monetary Fund. The data analysis in this work was performed while Luigi Guiso was at the Bank of Italy.

[1] See, among others, Diamond (1991a), Rajan (1992), Berglöf and Von Thadden (1994), Detragiache (1994), and Houston and James (1996).

[2] Sharpe (1990), Diamond (1991a), and Rajan (1992) present theoretical models. Lummer and McConnell (1989) show evidence supporting the idea that banks possess information not shared by other agents. Petersen and Rajan (1994, 1995) and Berger and Udell (1995) study various features of bank-firm relationships using small business data for the United States.

Table I

Number of Bank Relations for Small Businesses in Italy and in the United States

The data refer to firms with 500 or fewer employees. For the United States, they come from the 1988 wave of the National Survey of Small Business Finances. The number of relationships is the number of credit institutions (commercial banks, savings and loans associations, savings banks, credit unions, or mortgage banks) lending to the firm. For Italy, data are from the 1994 Survey of Manufacturing Firms and from the Credit Register. The number of relationships is the number of banks extending credit lines to the firm.

	United States	Italy
Mode	1	3
Median	2	5
75th percentile	2	8
Share of firms with one relationship	44.5%	11.0%

that firms borrow from one bank only, but this does not necessarily conform with the data. Table I reports some simple statistics on the number of bank relationships for samples of small firms in the United States and Italy. Single banking appears to be quite prevalent in the United States, but even there the median number of relationships is two, and 55.5 percent of firms have more than one bank. More striking is the case of Italy, where 89 percent of firms rely on multiple banking, the median number of relationships is five, and the 75th percentile is eight compared to two in the United States.

This paper addresses the questions of why firms, even rather small ones, need more than one bank and, more generally, what explains the cross-sectional variation in the number of banking relationships. A number of considerations suggest that multiple banking should be costly: dealing with more than one bank may involve significant transaction costs. Screening and monitoring costs may be duplicated, or, if banks free ride on the others' efforts, too little screening or monitoring may result. Also, debt renegotiation is likely to be more complex when many creditors are involved. To outweigh these costs, multiple banking must offer substantial benefits to firms or to banks. In the case of large firms, multiple banking may be explained by the desire of banks to diversify firm-specific credit risk. However, as shown in Table I, multiple banking is quite prevalent among small firms, especially in Italy.

One potential explanation for the benefits of multiple banking is suggested by Von Thadden (1994): After it acquires private information about the quality of the borrower, a relationship bank may be able to use this information to extract rents. Ex post rent extraction by the creditor, in turn, may distort entrepreneurial incentives ex ante, and lead to a suboptimal choice of investment projects. In these circumstances, establishing a second relationship can restore competition among lenders and limit ex post rent extraction.[3] In Von Thadden's model, *two* banking relationships are sufficient for the entrepreneur to capture all the rents, so why firms have *more than two banks* remains unexplained. In the model of Bolton and Scharfstein (1996), increasing the number of creditors complicates debt renegotiation, but the loss of ex post efficiency may be beneficial ex ante, as it limits incentives for strategic default. In equilibrium, firms with high credit quality are better off with two creditors rather than one, and so are firms whose assets are difficult to redeploy. Creditors, however, need not be relationship banks, and results for the case of more than two creditors are not derived.

In this paper, we present a theoretical model in which multiple banking can reduce "liquidity risk" in the sense of Diamond (1991b), namely the risk that a profitable project will have to be prematurely liquidated. By financing the early phases of a project, banks acquire private information about the continuation value of the venture. At the refinancing stage, the firm may need to borrow from nonrelationship banks because unexpected liquidity problems may make relationship banks unable to roll over their initial loans. In doing so, however, the firm faces an adverse selection problem, as noninformed banks will suspect that the project is a lemon.[4] For some parameter values, adverse selection is so severe that the firm is unable to refinance the project outside of the relationship. Under these circumstances, establishing multiple relationships serves to increase the probability that at least one informed bank will be able to refinance the project, thus reducing the likelihood of early liquidation. On the other hand, when adverse selection is mild enough that firms expect to be able to refinance from nonrelationship banks (albeit at a premium), multiple banking offers no benefits to the borrower, and single banking prevails.

The theoretical model has testable empirical implications concerning the determinants of the regime (single versus multiple banking) and concerning the optimal number of relationships for firms in the multiple banking regime. We explore the empirical validity of these implications using a cross section of small and medium-sized Italian manufacturing firms. The Italian data are especially useful because the number of relationships varies considerably across firms; moreover, by merging the data set with firm-level credit register data and with bank balance sheet data, we are able to obtain

[3] An alternative solution would be to resort to nonbank finance (Rajan, (1992)), and the greater availability of nonbank credit in the United States relative to Italy may be one of the factors explaining the greater prevalence of single banking in that country. Small firms, however, have little access to nonbank finance in the United States as well as in Italy. Houston and James (1996) study the firm characteristics associated with different types of debt finance, including multiple versus single banking, for a sample of large, publicly traded U.S. firms. They find that firm size and age are positively correlated with multiple banking, and so is the share of bank debt in total debt. The latter result is consistent with the hypothesis that for large firms multiple banking may perform the same role as nonbank finance.

[4] A similar adverse selection problem arises in Rajan (1992). In that model, no pure strategy equilibrium with refinancing outside of the relationship exists.

detailed information on the characteristics of each firm, on the firm's relationship with banks, and on the attributes of the banks lending to each individual firm. The predictions of the theoretical model suggest a testing strategy based on a two-stage estimation procedure: in the first stage, the choice of regime (single versus multiple banking) is estimated using a probit regression, and in the second stage we estimate the optimal number of relationships for the subsample of firms with more than one relationship taking into account the endogeneity of the regime. In both regressions we extend the set of explanatory variables beyond those suggested by our model to account for factors that could play a role based on intuitive considerations or alternative theories. Our findings are broadly consistent with the predictions of the theoretical model, especially concerning the choice of regime.

The paper is organized as follows. Section I presents the theoretical model and derives its empirical implications. Section II tests the predictions of the model using a sample of matched bank-firm data. Section III concludes.

I. The Theoretical Model

A. The Basic Setup

Consider a risk-neutral firm (or entrepreneur) that has no financial resources and wants to implement a project. The project starts at date 0 with investment cost I_0. It yields a return (cash flow) K at date 2 with probability p, and a zero return with the complementary probability, $1 - p$. At an intermediate date (date 1) the project needs a second allotment of funds, I_1, in order to be completed. If the cost I_1 is not paid, then the project must be liquidated at date 1, yielding a liquidation value of zero. The values I_0, I_1, and K are publicly known at date 0; that of p is not known by market participants, but it is common knowledge that it is the realization of a random variable with distribution $F(p)$, density $f(p)$, and support $[p_l, 1]$, with $p_l > 0$. At date 1 the realization of p is observed by the entrepreneur and by outside investors who financed the first stage of the project, but not by other parties. This assumption captures the idea that banks establishing a relationship with a firm acquire information that is not available to other banks. To avoid dealing with strategic interaction among banks at the monitoring stage, it will be assumed that information is acquired at no cost.[5] To keep the analysis simple, we consider only situations in which continuing the investment project at $t = 1$ is more efficient than liquidating it with probability one:

[5] If information were acquired through costly monitoring, and each bank could observe the lending terms offered by other banks, multiple banking would dilute the incentives of each single bank to monitor, and too little monitoring would take place in equilibrium. As a result, inside information would become less precise, mitigating the adverse selection problem and making refinancing outside of the relationship easier. To the extent that it results in less information acquisition, multiple banking may be similar to nonmonitored sources of external finance.

ASSUMPTION 1: $p_1K - I_1 \geq 0$.

Because the entrepreneur has no internal financial resources, he must obtain the amounts necessary for investment from outside investors. The latter consist of a large number of ex ante identical, perfectly competitive, risk-neutral banks, whose opportunity cost of funds is zero at both dates. Each bank incurs a fixed cost c for processing the loan, so multiple banking relationships increase transaction costs. Because we think of our model as especially suited for small and medium-sized firms, we neglect the possibility of nonbank external financing. Furthermore, we consider only one-period loan contracts.[6] Thus, to undertake the investment project the entrepreneur must borrow the amount I_0 at $t = 0$, and at $t = 1$ he must repay the initial loan and invest an additional I_1.

Imperfect loan enforcement is another key feature of the model, as it provides one of the frictions needed for multiple banking relationships to be beneficial. Because loan recovery through legal action is costly, banks can extract only a fraction $v < 1$ of the value of the firm from the borrower. Thus, at the final date, $t = 2$, creditors will receive a payment of no more than vK even if the contractual payment due exceeds that amount.[7] If loan enforcement is more difficult when there are many creditors, then the parameter v may increase with the number of relationships. To keep things simple, however, we will ignore this possible effect.

A second friction that plays a key role in the model is that banks may not be able to refinance a firm even if they believe that the project is profitable. This may occur when banks must reduce the size of their loan portfolios as the result of an unexpected loss of liquidity due to a flight of deposits. Alternatively, a bank's loan portfolio may become too risky relative to regulatory requirements or to the bank's own standards, and some risky projects must be dropped from the portfolio. We will refer to such circumstances as a "liquidity crisis" for the bank, and to a bank that does not experience such a crisis as an "active" bank. Liquidity crises occur with probability $\varepsilon > 0$ and are independent events across banks.[8] Thus, if a firm borrows from n banks at $t = 0$, the probability that at least one of those banks is active at $t = 1$ is $1 - \varepsilon^n$. Hence, increasing the number of banking relationships at $t = 0$ reduces the probability that all relationship banks will experience a crisis at the refinancing stage, $t = 1$.

[6] On the implications of introducing two-period loan contracts, see Section I.F below.

[7] This could be the equilibrium of a prebankruptcy debt renegotiation game in which creditors enforce their claims through the threat of bankruptcy and bankruptcy has deadweight costs. Another interpretation of the parameter v is that it reflects the control rents of the entrepreneur: The continuation value of the firm may depend on its current owner remaining at the helm, so the threat of expropriation through legal action following a default has limited effectiveness.

[8] The results hold if liquidity shocks are less than perfectly positively correlated across banks.

The number of banks participating in the first round of lending is common knowledge, but the occurrence of a liquidity crisis at any of the banks is not observable to other banks.[9] These assumptions imply that, if at $t = 1$ the firm attempts to refinance from a nonrelationship bank, other banks do not know if it is because all relationship banks are experiencing a liquidity crisis or because the project is bad (the realization of p is low).

B. The Refinancing Decision at $t = 1$

At the intermediate date $t = 1$, the firm must borrow to repay the first-period loan and to finance the second round of investment. Let D denote the (yet to be determined) interest and principal payment due on the first-period loan. At $t = 1$, borrowing from more than one bank is always suboptimal, as it increases transaction costs and has no possible benefits, so in equilibrium the firm refinances from one bank only, and transaction costs in equilibrium are always c. To economize on notation, I_1 denotes the amount of additional investment including the transaction cost c, so the amount that the firm must borrow at $t = 1$ is $D + I_1$. The firm can refinance from informed or uninformed banks. If the former option is chosen and if more than one informed bank is active, Bertrand competition ensures that the firm captures all the rents. If, on the other hand, only one informed bank is active, the firm and the bank are in a situation of bilateral monopoly, and the rent allocation is determined through bargaining. To simplify the analysis, however, we will assume that in equilibrium the borrower will capture all the rents even if only one informed bank is active at this stage. Accordingly, if R^i denotes the interest factor charged by an informed bank in equilibrium, the expected zero profit condition for informed banks is the following:

$$p \min[R^i(D + I_1), vK] - (D + I_1) = 0. \tag{1}$$

(Recall that because of imperfect enforcement, banks are unable to extract a repayment of more than vK from the firm.) It follows from equation (1) that if $p < (D + I_1)/vK$, then informed banks are not willing to refinance the project. On the other hand, if $p \geq (D + I_1)/vk$ then informed banks refinance the firm at the break-even interest factor $R^i = 1/p$. The term $vK/(D + I_1)$ is the maximum enforceable interest factor, and will be denoted by R^*. Consider now the strategy of an uninformed bank. If it offers an interest factor R, then the firm will take the loan in two circumstances: when no informed bank is active, an event with probability ε^n, or when at

[9] The assumption that banks know the number of creditors lending to a particular firm is quite realistic: in countries such as Italy or France this information is available at no cost through public credit registers. In countries such as the United States, where public credit registers do not exist, private credit bureaus provide the information at a small cost. Also, liquidity problems at other banks may be difficult to detect because banks may want to hide their problems from potential competitors and from depositors and borrowers.

least one informed bank is active and the interest rate offered by the uninformed bank is more favorable, namely if $R < R^i = 1/p$. The second event has probability $(1 - \varepsilon^n)F(1/R)$. Accordingly, the probability that a loan from an uniformed bank carrying an interest factor R is accepted is

$$\varepsilon^n + (1 - \varepsilon^n)F(1/R), \tag{2}$$

and the expected probability of success of the investment project conditional on refinancing from an uninformed lender is

$$E^u(p|R) = \frac{\varepsilon^n E(p) + (1 - \varepsilon^n)\int_{p_t}^{1/R} pf(p)\,dp}{\varepsilon^n + (1 - \varepsilon^n)F(1/R)}, \tag{3}$$

where $E(p)$ denotes the unconditional expected value of p. Since this expression exceeds $E(p)$, refinancing from uninformed banks reveals unfavorable information about the quality of the project. The break-even interest factor for uninformed lenders is

$$R^u = \frac{1}{E^u(p|R^u)}. \tag{4}$$

If R^u falls short of the highest enforceable interest factor R^*, then uninformed banks are willing to refinance the project at the intermediate date, and vice versa. Lemma 1 shows that a necessary and sufficient condition for uninformed bank refinancing is that expected profits be nonnegative at the maximum enforceable interest rate, namely that

$$R^*(D + I_1)E^u(p|R^*) - (D + I_1) > 0. \tag{5}$$

LEMMA 1: *Condition (5) is a necessary and sufficient condition for uninformed lenders to refinance the project at $t = 1$.*

Proof: See Appendix A.

The next step is to examine the borrowing decision at $t = 0$.

C. *An Equilibrium with Only One Banking Relationship*

If condition (5) holds, then the firm expects to be able to refinance the project with probability one at $t = 1$, and the first-period loan is always repaid. In equilibrium, then, the initial loan will carry the safe rate of interest, which is zero, so the amount borrowed at $t = 0$ and the amount to be repaid at $t = 1$ are the same. Furthermore, because the loan must cover transaction costs as well as the initial investment, this amount is $D = I_0 + nc$. The present discounted value of the firm's expected profits at $t = 0$ is then

$$\pi(n) \equiv E(p)K - (I_0 + nc + I_1)\left(R^u \varepsilon^n E(p) + (1 - \varepsilon^n)\left[R^u \int_{p_l}^{1/R^u} p\, dF(p)\right]\right). \tag{6}$$

Using equations (3) and (4), this expression becomes

$$\pi(n) = E(p)K - (I_0 + I_1 + nc). \tag{7}$$

Thus, if nonrelationship finance at date 1 is available, then there are no benefits to multiple banking: choosing $n > 1$ simply increases transaction costs with no offsetting benefits. This result is stated in Proposition 1.

PROPOSITION 1: *If condition (5) holds for $n = 1$ and for $D = I_0 + c$, so that refinancing from uninformed banks is possible, then in equilibrium the firm borrows from one bank only.*

D. An Equilibrium with Multiple Banking

If at the initial date the firm anticipates not to be able to refinance the project from uninformed creditors, then with probability ε^n the project will have to be liquidated prematurely, and with the complementary probability the project will be refinanced by an informed bank. In Section B it was shown that if $p > 1/R^*$ then an informed, active bank is willing to lend because the break-even interest factor $R^i = 1/p$ falls short of the maximum enforceable interest factor R^*. If $p < 1/R^*$, then refinancing will yield an expected loss whatever the interest rate charged. Lending, however, may still be in the best interest of an informed bank because such a bank is already exposed toward the firm and an early liquidation would lead to the loss of the first-period loan. Thus, creditors who are already exposed must choose the lesser of two evils: refinancing the project ("throwing good money after bad"), which involves disbursing I_1 in exchange for an expected repayment of pvK, or offering no refinancing and losing the initial investment.[10] Clearly, when $pvK > I_1$ (or $p > I_1/vK$), refinancing is in the collective interest of relationship creditors, but when $pvK < I_1$, liquidation is preferable. Although banks compete for the privilege of lending for realizations of p such that new lending is profitable ($p > 1/R^*$), they would be more than happy to let another bank shoulder the burden of refinancing at a loss when

[10] If informed banks know that the firm can access uninformed finance, then they will not extend new loans when $p < 1/R^*$.

$1/R^* < p < I_1/vK$. The game that relationship banks play in these circumstances may have interesting aspects but it is beyond the scope of this paper, so we will simply assume that in such states of the world each creditor contributes to the new loan in equal proportion, and that the indebted firm captures all the rents.[11]

To sum up, when condition (5) fails to hold, the project is refinanced at date 1 if two conditions are satisfied: at least one relationship bank is active and the probability of success of the project is $p \geq I_1/vK$. When these conditions obtain, the expected payoff to creditors who refinance the project is $p[(1/p)(I_1 + D)] - I_1 = D$ if $p > 1/R^*$, and zero otherwise. With these results in hand, let us now turn to the borrowing decision at $t = 0$. If the firm borrows from n creditors, and the loan is shared in equal proportion, the expected zero-profit condition for each bank is

$$(1 - \varepsilon^n)(1 - F(1/R^*))(D/n) - [(I_0/n) + c] = 0; \tag{8}$$

hence,

$$I_0 + cn = (1 - \varepsilon^n)(1 - F(1/R^*))D. \tag{9}$$

The present discounted value of the firm's expected profits at $t = 0$ is then

$$\pi(n) = (1 - \varepsilon^n)\left(\int_{I_1/vK}^{1} (pK - I_1)f(p)\,dp - (1 - F(1/R^*))D\right). \tag{10}$$

Using equation (8) and rearranging,

$$\pi(n) = (1 - \varepsilon^n)\left(\int_{I_1/vK}^{1} (pK - I_1)f(p)\,dp\right) - (I_0 + cn). \tag{11}$$

This expression shows that, if the parameters are such that uninformed banks are unwilling to refinance the project, then increasing the number of initial banking relationships affects expected profits in two ways: it lowers profits because it increases transaction costs, and it increases profits because it makes it more likely that the project will be completed. Maximizing expected profits by choice of n (and neglecting the fact that n must be an integer) yields the result stated in Proposition 2.

[11] Banks are assumed to participate in the bailout even if they are experiencing a liquidity crisis. This may be possible if, for instance, other active banks extend loans to banks with liquidity problems to allow them to assume their share of the burden.

Proposition 2: *If uninformed banks do not refinance at $t = 1$, then the optimal number of banking relationships is*

$$n^* = \frac{\ln c - \ln(-Z \ln \varepsilon)}{\ln \varepsilon} \tag{12}$$

where

$$Z \equiv \int_{I_1/vK}^{1} (pK - 1_1) f(p)\, dp. \tag{13}$$

Thus, the optimal number of relationships is generally greater than one, although it could be one for some values of the parameters.[12]

To summarize, if parameters are such that refinancing from informed banks is possible, then the number of relationships is trivially one, as multiple banking has no benefits. Conversely, if refinancing is not expected from non-relationship banks, then the optimal number of relationships is n^*. In the next subsection we briefly discuss what would change in the model if long-term loan contracts were allowed.

E. Long-Term Contracts

The results obtained so far rely on the assumption that firms finance the project with two short-term loans. However, it could be argued that firms could insure against the premature liquidation of the project by using long-term debt. To analyze this possibility, suppose that at $t = 0$ the firms borrows I_0 and also negotiates to obtain a second allotment of funds I_1 at the interim date. The repayment on both installments will take place at $t = 2$. What happens now if the bank extending the two-period loan experiences liquidity problems at $t = 1$? One possibility is that it will try to reduce its portfolio of risky loans by selling the loan contract to another bank.[13] If the bank sells the loan to another informed bank, then presumably the loan could be sold at its fair price. But if the prospective buyer is an uninformed bank, the price will reflect the same adverse selection problem that affects refinancing. For some values of the parameters, this problem may be so severe that uninformed banks refuse to purchase the loan. If the cost of keeping the loan while experiencing a liquidity crisis is substantial, then banks (and firms, to whom the cost is passed on) may be better off using multiple banking relationships.

[12] For n^* to be positive it must be $c < -Z \ln \varepsilon$, which we assume to hold.

[13] If long-term debt is callable, then the analysis conducted for short-term debt applies. Interestingly, the general condition for loan contracts suggested to its members by the Italian Bankers Association is that "the bank has faculty to recede, limit, reduce, or suspend the line of credit at any moment, even through verbal communication and even when the loan was extended without an explicit temporal limit."

F. Empirical Implications

With the purpose of deriving testable empirical implications for the theoretical model, in this section we analyze how the size of the single banking region and the optimal number of banking relationships in the multiple banking region depend on the parameters of the model.[14] The single banking region is defined by inequality (5) computed for $n = 1$ and $D = I_0 + c$, so changes in the parameters that increase the left-hand side of the inequality tend to make single banking more likely. Intuitively, the two main factors affecting the size of the left-hand side of inequality (5) are the severity of the adverse selection problem, which affects the size of the "lemons premium" that uninformed banks charge to refinance the firm, and the efficiency of the loan enforcement mechanism, which determines whether high interest rates are enforceable. The efficiency of the loan enforcement mechanism is measured by the parameter v, so the higher v is, the more likely is single banking. The severity of the adverse selection problem, on the other hand, is a decreasing function of ε, the parameter measuring bank fragility, because when ε is high, it is less likely that the firm will seek to refinance from uninformed banks for strategic reasons. Thus, a higher value of ε makes single banking more likely. Furthermore, other things being equal, uninformed banks are more likely to offer refinancing the more profitable is the project. Proposition 3 formalizes this intuition.

PROPOSITION 3: *The equilibrium with single banking described in Proposition 1 obtains when the probability of a liquidity crisis for banks is high, when the loan enforcement mechanism is highly efficient, and when the investment project is very profitable.*

Proof: See Appendix A.

Thus, firms that borrow from banks with a low probability of liquidity crisis, that operate in an environment with inefficient loan enforcement, and whose projects are not very profitable are more likely to be in the multiple banking region. In this region, an increase in n increases the firm's expected profits because it increases the probability that at least one informed bank is active $(1 - \varepsilon^n)$. The larger ε is, the larger the effect of an increase in n on such probability, and the larger the optimal number of relationships. Furthermore, the likelihood of being refinanced depends also on whether the realization of p is large enough that informed, active banks are willing to lend $(p > I_1/vK)$. More efficient loan enforcement makes this more likely, thereby increasing the benefits of expanding the number of relationships. Thus, the optimal number of relationships is increasing in v.

[14] Comparative statics with respect to the transaction cost parameter c are not discussed because in the empirical section we have no direct proxy for this parameter.

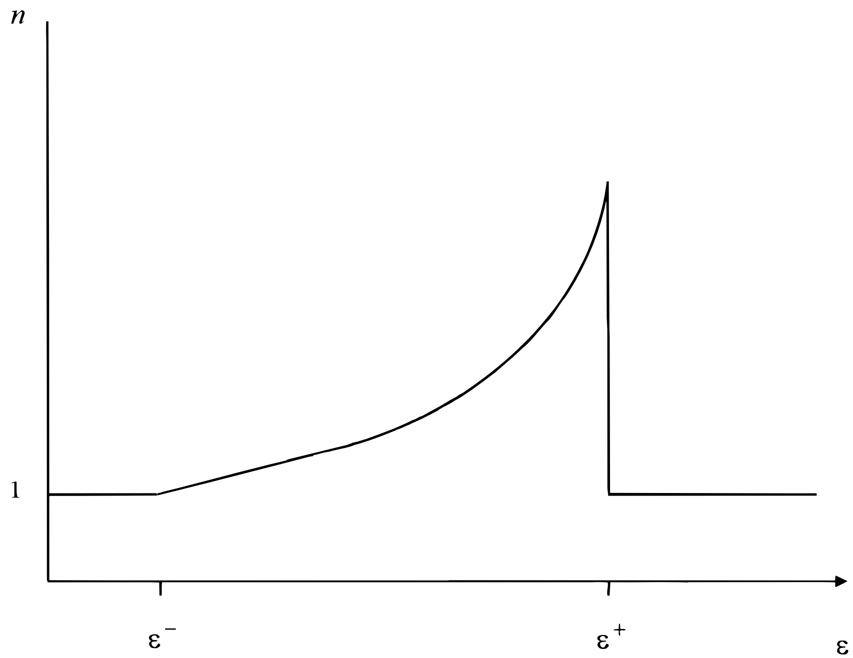


Figure 1. The optimal number of banking relationships. The figure shows the optimal number of relations (n) as a function of the value of ε, the probability of a liquidity crisis. ε^- and ε^+ are the thresholds for ε that define the multiple banking area. From $\varepsilon^- < \varepsilon < \varepsilon^+$ firms choose to borrow from multiple banks; for very low values of ε ($\varepsilon < \varepsilon^-$) or very high ($\varepsilon > \varepsilon^-$) a regime of single banking prevails.

Finally, the more profitable the project, the larger the loss from premature liquidation, and the more beneficial is multiple banking. These results are derived in the following proposition.

PROPOSITION 4: *In the multiple banking region, the optimal number of relationships is increasing in the efficiency of the loan enforcement mechanism and in the profitability of the second stage of the investment project. For realistic values of the parameters, it is also increasing in the probability of a liquidity crisis for banks.*

Proof: See Appendix A.

Propositions 3 and 4 imply a nonmonotonicity in the relationship between the optimal number of banks and the three parameters: large values of the parameters tend to decrease the likelihood that a firm will be in the multiple banking region, but, conditional on being in the multiple banking region, large values of the parameters should be associated with a larger number of banks. Figure 1 illustrates the relationship between the optimal number of relationships and the fragility parameter ε. In the figure, ε^+ is the value of ε for which condition (5) holds with equality. Thus, for $\varepsilon > \varepsilon^+$ the firm is in the single banking region, and the optimal number of relationships is one. For $\varepsilon < \varepsilon^+$ the firm is in the multiple banking region, and n^* is increasing in ε. For very low values of ε ($\varepsilon < \varepsilon^-$) the benefits of multiple banking are so low that choosing $n = 1$ is optimal even if the firm is in the multiple banking region. Similar figures could be drawn for the other parameters.

To gain a better understanding of how the number of relationships changes with the parameters of the model, we carried out some simulations using alternative values for the parameters. The results are reported in Table II for low and high transaction costs. The single banking outcome is denoted by "*S*" and it occurs for values of the fragility parameter falling in the region defined by inequality (5). When multiple banking prevails, the number of banks increases with the fragility parameter ε, as expected. When transaction costs are low and banks are relatively fragile (Panel A, last column), the optimal number of relationships can be as high as 10. The simulations indicate that changes in the enforcement parameter v have little effect on n^* in the multiple banking region. For instance, as shown in Panel B, increasing v from 0.50 to 0.90 while keeping $\varepsilon = 0.20$, $K = 3$, $I_0 = I_1 = 0.33$, leaves the number of relationships unchanged at three. Similar results apply to changes in the profitability parameters (a change in profitability is obtained by changing the investment requirement while keeping K constant). On the other hand, the value of ε^+, which marks the boundary between the single and the multiple banking region, is very sensitive to both v and to the profitability of the project. In the next to last column of Panel A one can observe how a switch in regime occurs when v is increased from 0.6 to 0.8. Similarly, in the second row a switch from multiple to single banking occurs when ε is increased from 0.3 to 0.5. To summarize, the simulations show that enforcement and profitability should have a strong impact on whether the firm has one or many relationships, but a small effect on the number of relationships for firms in the multiple banking regime. Bank fragility, on the other hand, should matter both for the choice of regime and for the number of relationships.

II. Testing the Theory

The theoretical model developed in the previous section suggests several firm and bank characteristics that should affect the number of banking relationships. In this section, we test the predictions of the model for a cross section of Italian firms. There are three reasons for relying on Italian data. First, among Italian firms there is considerable variation in the number of relationships, which can be as large as 30, even among relatively small businesses. Second, the efficiency of the judicial system varies significantly across regions and this allows us to use a natural proxy for the parameter v, namely the share of loans recovered through the judicial system following default.

Table II

Regions of Multiple Banking and Optimal Number of Relations for Various Parameter Values

The table reports the optimal number of relationships for various values of the parameters of the theoretical model. The single banking outcome is denoted by S. Also reported are the values of the parameter ε (denoted by ε^- and ε^+) marking the boundaries of the region in which the optimal number of relationships is greater than one. ε is the probability that a bank is hit by a liquidity shock. The parameter v is the maximum fraction of the cash flow that creditors can extract from the firm. c is the cost of setting up a relationship. I_0 and I_1 measure the investment costs of the project, and K is cash flow from the project if it succeeds. Given K, higher values of I_0 and/or I_1 imply a lower profitability of the project. All simulations assume that the probability of success of the project p is distributed uniformly in the interval (0.1, 1).

Parameter Values		Boundaries of the Multiple Banking Region		Optimal Number of Relations for Different Values of the Probability of a Liquidity Shock (ε)		
Profitability	Enforcement	Lower Bound (ε^-)	Upper Bound (ε^+)	$\varepsilon = 0.2$	$\varepsilon = 0.3$	$\varepsilon = 0.5$
Panel A: Low Cost of Setting Up a Relationship $c = 1/1000$						
Low profitability	$v = 0.5$	0.013	0.51	5	6	10
($K = 3$, $I_0 = I_1 = 0.2$)	$v = 0.6$	0.013	0.34	5	6	S
	$v = 0.8$	0.013	0.17	S	S	S
	$v = 0.9$	0.013	0.11	S	S	S
High profitability	$v = 0.5$	0.016	1.00	5	6	10
($K = 3$; $I_0 = I_1 = 0.33$)	$v = 0.6$	0.016	1.00	5	6	10
	$v = 0.8$	0.018	0.54	5	6	10
	$v = 0.9$	0.018	0.43	5	6	S
Panel B: High Cost of Setting Up a Relationship $c = 1/100$						
Low profitability	$v = 0.5$	0.013	0.53	3	4	6
($K = 3$, $I_0 = I_1 = 0.2$)	$v = 0.6$	0.013	0.36	3	4	S
	$v = 0.8$	0.013	0.18	S	S	S
	$v = 0.9$	0.013	0.13	S	S	S
High profitability	$v = 0.5$	0.054	1.00	3	4	6
($K = 3$; $I_0 = I_1 = 0.33$)	$v = 0.6$	0.054	1.00	3	4	6
	$v = 0.8$	0.055	0.56	3	4	6
	$v = 0.9$	0.055	0.43	3	4	S

Finally, we have obtained data that allow us to identify all the banks lending to each firm in the sample, a crucial feature to construct measures of shocks to bank liquidity.

A. Data Sources and Description

To construct the sample, we rely on three different data sources, which are described in detail in Appendix B. The first is the 1994 wave of the Survey of Manufacturing Firms (SMF) conducted by Mediocredito Centrale (an investment bank) on a random sample of more than 4,000 small and medium-sized manufacturing firms. The SMF reports various characteristics of the firms, but has no information on the firms' sources of funding or on bank-firm relationships. To obtain this information, we merged the SMF with data from the Credit Register (CR) database. The CR database reports, for each firm and loan category, the amount of credit granted by each relationship bank together with the amount utilized. The merged data set covers 1,849 firms and contains a full description of each firm and its indebtedness with all the banks in the system. It is our reference sample. A firm is considered to have a relationship with a bank if it borrows from the bank under at least one of the six loan categories reported in the data base.[15]

The third and final step was to obtain information on the characteristics of the banks lending to each individual firm. This was achieved by merging the data set with the Bank Balance Sheet database (BB), which reports for each bank in the system and for various years (at a quarterly frequency) information on balance sheet items and other bank characteristics.

B. The Testing Strategy

According to the theoretical model, firms can be in one of two regimes depending on parameter values: in the first regime, they expect to be able to access nonrelationship banks to refinance investment. Firms in this regime always find it optimal to borrow from one bank only, as multiple banking has costs but no benefits. In the second regime, loans from nonrelationship banks are not available at the refinancing stage, and more relationships reduce the probability of premature liquidation of the investment project.

These features of the theoretical model suggest a two-stage empirical test: in a first stage, the implications of the model concerning the choice of regime are tested, and in the second stage the number of relationships is estimated conditional on the firm being in the multiple banking regime. The first-stage estimation is performed using a probit model, in which the independent variable is a dummy equal to one for firms with only one relationship

[15] The loan categories are: commercial paper discounted by the bank, lines of credit, export loans, collateralized loans, medium-term loans, and long-term loans. All firms in our sample have at least one credit line; 1,393 discount commercial paper and 146 financial paper. Export loans are used by 998 firms; few firms have medium-term and long-term loans and collateralized loans.

and zero for firms with multiple relations. The explanatory variables are proxies for firm and bank characteristics that impinge on the size of the single-banking region as stated in Proposition 3 above.

To make the empirical specification richer, we also include variables capturing aspects suggested by alternative theories or likely to be important in practice. The second-stage regression estimates the number of banking relationships in the multiple banking region. Though the number of relationships is a discrete variable, for simplicity we treat it as a continuous variable, and use an OLS estimator. Since the number of relations has a large range of variation (from two to more than 30) this should not constitute a source of significant bias. To correct for selection bias, we insert the Mills ratio computed from the first-stage probit estimate in the OLS regression.

C. The Explanatory Variables

As a proxy for the enforcement parameter v we use the average share of credit recovered through the judicial system by banks operating in the same region as the firm. This share has considerable geographical variation due to differences in the efficiency of the judicial system.[16] The parameter v may also be affected by firm characteristics that are associated with larger entrepreneurial control rents, and therefore strengthen the bargaining power of management/shareholders relative to creditors in workout negotiations. Proxies for control rents included in the regression are the share of the firm owned by the largest shareholder and the correlation between the firms' output and the output of firms belonging to the same industry (see Appendix B for details on the construction of all the variables). The higher this correlation, the more difficult it should be for creditors to sell off the assets of a liquidated firm, and the larger should be the entrepreneurial rents (Shleifer and Vishny (1992)).

Additional proxies for entrepreneurial control rents are various measures of research and development and innovation intensity.[17] A negative correlation between the extent of entrepreneurial control rents and the probability of single banking also supports the hypothesis that multiple banking serves to reduce rent appropriation by banks (Von Thadden (1994)).

The other explanatory variables suggested by the model are the profitability of the investment project (proxied by the profit-to-sales ratio) and bank fragility, namely the propensity of banks to experience liquidity problems that force them to cut back their loan portfolios. To capture bank fragility, we rely on two indicators. The first is a proxy for idiosyncratic liquidity shocks to each bank based on observed changes in the ratio of liquid funds

[16] For instance, the average length of time to repossess collateral is about three years in Valle d'Aosta, a Northern region, and about seven years in the Southern region of Sicily (Generale and Gobbi (1996)). Creditors must use the courts of the region where the firm is located to recover defaulted loans.

[17] Alternatively, R&D intensity may be associated with single banking if information leakages to competitors are more likely with multiple lenders (Yosha (1995)).

to assets (see Appendix B for details). The second proxy is a weighted average of the ratio of nonperforming loans to assets of each creditor bank. In this case, the hypothesis is that banks with a higher share of nonperforming loans are more prone to liquidity shocks.

Because many features of the economy in Italy exhibit sharp differences across regions, we introduce dummy variables for firms operating in Northern Italy and in Central Italy. In all the regressions we control for the size of the firm measured by the number of employees. Firm size is expected to affect the number of relationship for three reasons. First, large firms may have to rely on multiple banking to allow banks to diversify firm-specific credit risk. Second, firm "complexity" is likely to increase with size, and more complex firms may find it necessary to rely on several banks (e.g., because they need to serve plants located in different regions). Finally, if the cost of setting up a new relationship is in part fixed, then larger firms with larger borrowing requirements should be more inclined to rely on multiple banking.

Other explanatory variables included are the leverage of the firm, its age, and a dummy variable for group membership. Leverage may increase the probability of multiple banking, because—other things being equal—the probability of default is higher for more leveraged firms, and the adverse selection problem should be more severe. Also, higher leverage may mean lower unit costs of establishing a relationship. Firms belonging to a group may be better able to obtain financing from other members of the group when nonrelationship banks are unwilling to refinance, so we expect group membership to be associated with more reliance on single banking and with a smaller number of relationships. Older firms, to the extent that they are better known, may face less severe adverse selection problems when seeking nonrelationship finance, so they may be less likely to use multiple banking as well. Also, if building banking relationships takes time, in the multiple banking region age may be positively correlated with the number of relationships.

Finally, we introduce as an additional control the average size of the banks lending to each firm. Bank size may be related to bank fragility; as shown by Buchinsky and Yosha (1995), smaller banks are more likely to be exposed to liquidity shocks. Bank size should then have a negative effect on the probability of borrowing from a single bank and on the number of relationships in the multiple banking regime. Table III summarizes the explanatory variables included both in the probit and in the OLS regression and their expected sign.

D. Results

Table IV shows the results of the probit regression for the probability of single banking in column (1). The average share of loans recovered after default, proxying the parameter v, has a positive and significant coefficient as predicted by the model. Its effect is also economically relevant; raising the share of loans recovered from its minimum value of about 40 percent to

Table III

Explanatory Variables and Their Expected Sign

The table lists the explanatory variables used in the regressions and the expected signs of their coefficients according to the theoretical model. The notation (or zero) indicates that the variable has a negligible impact on the dependent variable according to simulation results. The share of defaulted loans recovered is the average share recovered through legal procedures by the banks operating in the firm's province. Liquidity shocks are defined as the average variability of the innovations to a measure of liquidity for the pool of banks lending to the firm. Nonperforming loans are scaled by loanable funds. Profitability is gross profits divided by sales. Leverage is the ratio of total bank debt to sales. Firm size is the number of employees. Firm age is the number of years since the firm was founded. For each firm the size of lending banks is the average value of loanable funds of its relationship banks. The comovement indicator is a measure of correlation of sales among firms in the same industry using a 44 industrial sector classification. See Appendix B for more detailed information on the construction of the variables.

Explanatory Variable	Probability of Single Banking	Number of Relationships in the Multiple Banking Region
Share of defaulted loan recovered	+	− (or zero)
Liquidity shocks	+	+
Nonperforming loans	+	+
Profitability	+	+
Firm size	−	+
Firm leverage	−	+
Firm age	+	+
Group membership	+	?
Average size of lending banks	−	−
Patents	−	+ (or zero)
R&D	−	+ (or zero)
Product innovation	−	+ (or zero)
Process innovation	−	+ (or zero)
Product and process innovation	−	+ (or zero)
Share of first owner	−	+ (or zero)
Industry comovement	−	+ (or zero)

85 percent[18] would increase the probability of single banking by nine percentage points. This is a remarkable effect, given that the unconditional probability of single banking is 10 percent. The variables measuring bank fragility have the expected signs and are statistically significant (although the ratio of nonperforming loans is significant only at the 10 percent confidence level), but the size of the effect is small: increasing the two variables by two standard deviations would increase the probability of single banking by only 1.4 and 1.3 percentage points, respectively. This is consistent with the simulation results in Table II. Profitability also increases the probability

[18] This figure is close to the recovery rate for structured loans in the United States (Asarnow and Edwards (1995)).

of single banking, as predicted by the theoretical model. Thus, the three variables suggested by the model seem to play a relevant role in determining whether a firm relies on single or on multiple banking.

For variables not explicitly modeled, the table shows that firm size has a negative and significant impact, confirming that larger firms are less likely to use single banking, and so does leverage, as expected. Also, it appears that firms borrowing from large banks are—ceteris paribus—less likely to rely on single banking than firms borrowing from small banks. This result is consistent with the view that bank size is a proxy for bank liquidity. Firm age, the group membership dummy, and the variables capturing entrepreneurial control rents are not statistically significant. Finally, of the five indicators of firm's attitude to innovate, all but one have a negative effect on the probability of single banking, as expected. All of the coefficients, however, have relatively high standard errors. As a robustness check, the probit regression was reestimated omitting all the variables that were not statistically significant in the previous specification; the size of the coefficient of the remaining explanatory variables and the significance levels remain virtually unchanged.

Table IV, column (3), reports the estimation results of the second-stage regression, in which the determinants of the number of relationships for firms with more than two banks are estimated. To avoid relying on the nonlinearity of the probit estimates to reach identification of the second-stage parameter, we assume that all the variables related to R&D and innovation adoption, though they affect the choice of regime, do not affect the number of relationships in the second stage. This restriction is justified by the simulation results in Table II, which show that proxies for v, though they affect the choice of regime, do not affect the number of relations in the multiple banking regime.[19] To correct for the selection bias, we insert as a regressor the Mill's ratio computed from the first-stage probit. The proxy for loan recovery is not significant and has the wrong sign; however, because the simulation results show that the number of relationships in the multiple banking regime is not sensitive to changes in this parameter, this result should not be taken as evidence against the theoretical model. In contrast, both variables capturing bank fragility are significant and have positive coefficients as predicted by the model. The average size of the lending banks has a negative and significant effect, and this finding, together with the negative coefficient of bank size in the probit equation, supports the interpretation of this variable as a proxy for bank liquidity. Firm profitability has the predicted positive sign but is not statistically significant. Finally, the number of relationships tends to increase with firm size and leverage, but the group membership dummy and the variables capturing entrepreneurial

[19] An alternative procedure to reach identification of the second-stage parameters would be to exclude the loan recovery variable v. Since this variable is not significant in the second-stage regression (see below), following this route would not change the results.

Table IV

Determinants of the Choice between Single and Multiple Banking and of the Optimal Number of Banking Relationships in the Multiple Banking Region

The table reports regression coefficients and associated t-statistics. The first two regressions estimate the probability that the firm borrows from only one bank using a probit model. In the last two regressions the dependent variable is the number of relationships for firms using more than one bank. The regressions are estimated using OLS corrected for selectivity bias using the Mill's ratio computed from the first-stage probit. The Mill's ratios of regressions (3) and (4) are computed from the probit regressions (1) and (2) respectively. The share of defaulted loans recovered is the average share recovered through legal procedures by the banks operating in the firm's province. Liquidity shocks are measured by the average variability of the innovations to a measure of liquidity for the pool of banks lending to the firm. Nonperforming loans are scaled by loanable funds. Profitability is gross profits divided by sales. Leverage is the ratio of beginning-of-period total bank debt to sales. Firm size is the number of employees. Firm age is the number of years since the firm was founded. The comovement indicator is a measure of correlation of sales among firms in the same industry using a 44 industrial sector classification. See Appendix B for more detailed information on the construction of the variables. The (0,1) notation means that the variable is a dummy equal to one if the firm has the specified characteristic. All regressions include two regional dummies for the firm location.

	Probability of Single Banking		Optimal Number of Relationships in the Multiple Banking Region	
Explanatory Variable	(1)	(2)	(3)	(4)
Share of defaulted loans recovered	0.016	0.018	−0.001	−0.015
	(2.175)	(2.454)	(−0.068)	(−0.971)
Liquidity shocks	0.504		2.141	2.510
	(2.333)		(2.081)	(2.479)
(Liquidity shocks)$^2 \mid Q > 10$		−3.841		
		(−5.629)		
(Liquidity shocks)$^2 \mid Q \leq 10$		1.657		
		(6.634)		
Nonperforming loans	5.589		26.016	9.759
	(1.750)		(2.476)	(1.346)
(Nonperforming loans)$^2 \mid Q \leq 10$		−17.465		
		(−0.822)		
(Nonperforming loans)$^2 \mid Q > 10$		15.558		
		(0.741)		
Profitability	2.164	1.757	0.834	0.053
	(3.196)	(2.608)	(1.009)	(0.067)
Firm size	−0.187	−0.135	0.961	1.133
	(−3.221)	(−2.241)	(7.276)	(9.581)
Firm leverage	−1.708	−1.576	3.968	4.739
	(−6.421)	(−5.805)	(6.228)	(8.792)
Firm age	0.002	0.002	0.0215	0.0189
	(0.902)	(0.831)	(4.479)	(4.026)
Group membership (0,1)	−0.017	−0.019	0.070	0.093
	(−0.163)	(−0.175)	(0.325)	(0.436)
Average size of lending banks	−0.248	−0.188	−0.336	−0.121
	(−6.400)	(−4.700)	(−2.039)	(−0.949)

Table IV—*Continued*

	Probability of Single Banking		Optimal Number of Relationships in the Multiple Banking Region	
Explanatory Variable	(1)	(2)	(3)	(4)
Patents (0,1)	−0.212	−0.212		
	(−1.197)	(−1.197)		
R&D (0,1)	−0.147	−0.137		
	(−1.300)	(−0.748)		
Product innovation (0,1)	−0.0429	−0.0039		
	(−0.278)	(−0.243)		
Process innovation (0,1)	−0.165	−0.162		
	(−1.313)	(−1.237)		
Product and process innovation (0,1)	0.117	0.109		
	(0.970)	(0.880)		
Share of first owner	−0.116	0.002	0.005	0.003
	(−1.210)	(1.431)	(1.370)	(0.973)
Industry comovement	−0.163	0.879	2.312	2.306
	(−0.910)	(1.368)	(1.712)	(1.737)
Constant	0.963	2.507	2.486	−0.640
	(1.371)	(2.759)	(1.125)	(−0.351)
Mill's ratio			−10.446	−7.506
			(−7.289)	(−10.01)
Number of observations	1,754	1,760	1,557	1,563

control rents are not significant. Overall, the results of the estimation are remarkably supportive of the predictions of the model, although the significance of variables such as firm size, leverage, and age suggests that there are factors other than those considered by our theory that are empirically relevant.

E. Extensions

The above estimates have one potential drawback: According to the theoretical model, all firms in the single-banking region of the parameter space (i.e., firms fulfilling condition (5)) have one relationship. However, there may also be firms belonging to the multiple banking region who choose to borrow from one bank only because the benefits from multiple banking are very low relative to the costs (see Section I.F). The approach described so far assumes that all firms with only one relationship in the sample belong to the former category, and to the extent that this assumption is not correct, the probit regression may be misspecified. Take for example the parameter ε, capturing bank fragility. The smaller is ε, the less likely the firm is to be in the single banking region (Proposition 3); however, if ε is very small, then the firm could be in the multiple banking region *and* have chosen $n^* = 1$. The same is true for the parameter v. In the case of v, however, the simulations in Table II indicate that firms in the multiple banking region choose $n^* = 1$ only for implausibly low values of v, so that in this case the misspecification problem can safely be ignored.

To take into account potential nonmonotonicities with respect to bank fragility in the choice of regime, we estimate an alternative specification of the probit equation where ε is replaced by two truncated variables. The first is equal to the square of ε if ε is less than or equal to its 10th percentile (our empirical threshold for ε^-); otherwise it is equal to the 10th percentile. The second truncated variable is equal to the square of ε if ε is greater than its 10th percentile, and equal to the 10th percentile otherwise. To support the prediction of the theoretical model that the relation between the probability of single banking and liquidity shocks has an inverse U-shape (decreasing for low values of ε and increasing for relatively high values) we should find a negative coefficient for the first truncated variable and a positive coefficient for the second truncated variable. The results of the probit estimation are reported in the column (2) of Table IV, and those of the corresponding second-stage regression are in the last column. Inserting the nonlinear terms in ε has little effect on the coefficients of the other variables in the probit, and nothing of substance changes in the second-stage regression.

The presence of the nonmonotonicities suggested by the theory is supported by the coefficient estimates when the liquidity shock is used as a proxy for ε: the coefficient of the first truncated variable is negative, that of the second is positive, and both are statistically significant. When the share of nonperforming loans is used as a proxy for liquidity shocks, the sign of the coefficients are as predicted but the standard errors are large.[20]

F. Evidence from Cross-Country Comparison

In the introduction, we reported data on bank-firm relationships in Italy and the United States, indicating that, although both U.S. and Italian firms make use of multiple banking, single banking tends to be much more prevalent in the United States. This evidence raises the question of whether the number of relationships depends not only on firm-level characteristics, but also on features of the financial system or the economy that may be different across countries. To properly address this question would require considerable data gathering efforts and is well beyond the scope of this paper. Nonetheless, we have been able to obtain data on the number of bank relationships and on the efficiency of the loan recovery mechanism for 11 European countries (including Italy). The data on the number of relationships come from a survey of corporate treasurers and cash managers of 1,129 firms in 20 European countries. This data set is analyzed in a recent study by Ongena and

[20] Results (not reported) are similar if we capture nonmonotonicities in the probit regression by inserting quadratic terms in addition to the linear terms for the liquidity shocks proxies.

Smith (1998).[21] Unfortunately, the number of observations per country tends to be small (e.g., there are only 25 French firms and 13 Polish firms), and it is unclear to what extent the sample is representative of the population. Thus, the evidence presented in this section should be regarded as tentative.

To measure the efficiency of loan recovery, we use data on collateral recovery costs and on time to recovery for mortgage loans from the European Mortgage Lenders Association. These figures are available for 11 countries. We find a negative correlation between the inefficiency of loan recovery and the incidence of single banking, as predicted by our model, with correlation coefficients of -0.67 when collateral recovery cost is used as a proxy of v and -0.46 percent when the time to recovery is used.

III. Concluding Remarks

Much of the theory of banking relationships implicitly assumes that firms have only one bank, but empirical evidence shows that this is not always the case. In particular, in a sample of small and medium-sized Italian firms we find that multiple banking is extremely widespread. To explore why firms may find it necessary to establish multiple relationships with banks, we have developed a theoretical model where multiple banking ensures a more stable supply of credit, and reduces the risk of premature liquidation of the investment project. The intuition is quite straightforward: If there is a positive probability that the relationship will be terminated because of reasons internal to the bank (e.g., liquidity problems), the firm may expect to have to resort to nonrelationship finance at some time in the future. But in accessing this source of credit, the firm faces an adverse selection problem, as outsiders have less information than relationship lenders and suspect that the firm may be a bad risk. When the adverse selection problem is so severe that noninformed finance is unavailable, then it may pay off to establish more than one relationship from the start.

The theoretical model predicts that firms whose banks are less fragile or whose creditors find it more difficult to enforce loan repayment should be more likely to resort to multiple banking. Conditional on adopting multiple banking, however, bank fragility should be associated with a larger number of relationships. The expected profitability of the project has effects similar to bank fragility. We test whether these predictions receive empirical support using firm level data for a sample of Italian small and medium-sized

[21] Ongena and Smith (1998) perform a number of interesting tests on this data set. For instance, they test whether the cross-country variation in the number of relationship is explained by such variables as the concentration of the banking industry, the development of equity and bond markets, the efficiency of the judicial system, the degree of protection of creditor rights, and the fragility of the banking system. Introduction of the latter three variables was inspired by the theory presented in this paper. However, the authors do not follow the two-stage estimation procedure required by our theoretical model; thus it is not clear that their findings can be related to the predictions of our model.

businesses. This data set contains detailed information on bank-firm relationships and the characteristics of the banks lending to each firm. The empirical results lend support to the theory, especially concerning the determinants of the choice between single and multiple banking.

We also offer some evidence from a cross section of European countries showing that the efficiency of the loan enforcement mechanism tends to be negatively correlated with the share of firms using a single banking relationship, as predicted by our theory. A comprehensive empirical analysis of the factors explaining the configuration of bank-firm ties across countries remains a topic for future research.

Appendix A. Proofs

Proof of Lemma 1: To prove that condition (5) is sufficient, notice that expected profits of uninformed lenders are negative for $R = 1$, and that they are a continuous function of R for $R \geq 1/p_l$. Thus, if the inequality holds there must be a value of R for which uninformed banks break even. To prove necessity, it is sufficient to show that expected profits are increasing in R; actually, all that is needed is that expected profits be increasing when nonnegative since negative expected profits for $R = 1$ and for $R = R^*$ would imply negative expected profits for all $R \in [1, R^*]$. To this end, rewrite equation (3) as

$$\frac{a(R)}{b(R)} - 1 = 0 \tag{A1}$$

where

$$a(R) = R\left[\varepsilon^n E(p) + (1 - \varepsilon^n)\int_{p_l}^{(1/R)} pf(p)\,dp\right]$$

and

$$b(R) = \varepsilon^n + (1 - \varepsilon^n)F(1/R).$$

Since $b(R) > 0$, the left-hand side of equation (A1) is increasing if and only if

$$a'(R) - b'(R)\left[\frac{a(R)}{b(R)}\right] > 0. \tag{A2}$$

Since

$$a'(R) = \varepsilon^n E(p) + (1 - \varepsilon^n)\int_{P_1}^{(1/R)} pf(p)\,dp - (1 - \varepsilon^n)f(1/R)(1/R)^2$$

and

$$b'(R) = -(1 - \varepsilon^n) f(1/R)(1/R)^2,$$

inequality (A2) can be rewritten as

$$\varepsilon^n E(p) + (1 - \varepsilon^n) \int_{p_l}^{(1/R)} p f(p)\, dp - (1 - \varepsilon^n) f(1/R)(1/R)^2 \left[1 - \frac{a(R)}{b(R)} \right] > 0. \tag{A3}$$

But when expected profits are nonnegative, $a(R)/b(R) \geq 1$, thus the condition above must hold.

Proof of Proposition 3: The region of single banking is defined by inequality (5) calculated for $n = 1$ and $D = I_0 + c$. Let $T(\varepsilon, v, K, I_0, I_1)$ be the left-hand side of inequality (5). To prove the proposition, we will show that the function T is increasing in its first three arguments and decreasing in the last two arguments (an increase in K for constant investment levels corresponds to an increase in the profitability of the project, and so does a decline in I_0 or I_1 for constant K). To this end, define $q = (I_0 + I_1 + c)/vK$. Then

$$\left. \frac{\partial T}{\partial \varepsilon} \right|_{n=1} = (1/q) \left[E(p) - \int_{p_l}^{q} p f(p)\, dp \right] - 1 + F(q). \tag{A4}$$

Hence,

$$\left. \frac{\partial T}{\partial \varepsilon} \right|_{n=1} = \int_{q}^{1} \left(\frac{p}{q} - 1 \right) f(p)\, dp, \tag{A5}$$

which is always positive. To prove that T is increasing in v and in the profitability of the project, it is sufficient to prove that T is decreasing in q. To see this, notice first that v, K, I_0, and I_1 enter T only through q, and that T is decreasing in v and K and increasing in I_0 and I_1. The derivative of T with respect to q is

$$\left. \frac{\partial T}{\partial q} \right|_{n=1} = (1/q^2) \left(q^2 (1 - \varepsilon) f(q) - \left[\varepsilon E p + (1 - \varepsilon) \int_{p_l}^{q} p f(p)\, dp \right] \right) - f(q)(1 - \varepsilon); \tag{A6}$$

hence,

$$\left.\frac{\partial T}{\partial q}\right|_{n=1} = -(1/q^2)\left(\varepsilon Ep + (1-\varepsilon)\int_{p_l}^{q} pf(q)\,dp\right). \tag{A7}$$

This expression is always negative.

Proof of Proposition 4: First, we will prove that n^* is increasing in the probability of a liquidity crisis, ε, at least for reasonable values of the parameters. Define $x = -\ln \varepsilon > 0$. Then, from equation (12)

$$n^* = \frac{-\ln c + \ln(Zx)}{x}, \tag{A8}$$

so that n^* is increasing in ε if and only if it is decreasing in x. Differentiating with respect to x,

$$\frac{\partial n^*}{\partial x} = \frac{1 - [\ln(Zx) - \ln c]}{x^2}. \tag{A9}$$

Rearranging, for n^* to be increasing in ε it must be

$$\ln Z > \ln c + 1 - \ln \varepsilon. \tag{A10}$$

This is a mild condition on the expected profitability of the second stage of the project. For instance, if transaction costs are as high as five percent and the probability of a liquidity crisis is as high as 30 percent, the condition is satisfied if the expected return is 11.3 percent. If transaction costs are one percent and ε is still 30 percent, the threshold value of Z falls to 2.3 percent. Lower values for the probability of a liquidity crisis would further reduce the lower bound on Z. Thus, for realistic values of the parameters the optimal number of relationships n^* is increasing in ε. As for the effects of v on n^*, note that

$$\frac{\partial n^*}{\partial v} = \left(-\frac{1}{Z\ln\varepsilon}\right)\left(\frac{\partial Z}{\partial v}\right) = \left(-\frac{1}{Z\ln\varepsilon}\right)\left(\frac{(I_1)^2}{(v)^2K}\right)\left(\frac{1-v}{v}\right)f\left(\frac{I_1}{vK}\right), \tag{A11}$$

which is always positive. Finally, the equation above also shows that an increase in Z, the expected continuation value of the project, increases n^*.

Appendix B

A. Data Sources

Three main sources have been used in the empirical analysis: the 1994 Survey of Manufacturing Firms (SMF), the Credit Register information (CR), and the Banks Balance Sheet database (BB). The SMF is conducted every three years on a sample of small and medium-sized manufacturing firms with at least 10 employees. The sample contains 4,431 firms. The information collected includes employment and its composition, investment effected and types of investments, R&D activities, location, ownership structure, industrial sector, year of foundation, capacity utilization, total sales, export sales and a number of balance sheet items for the survey year and the two previous years. For firms with fewer than 500 employees, the sample is stratified by gross product per employee in order to ensure representativeness. Also, detailed information on mergers, acquisitions, and breakups is available. For a subset of the firms in the sample, credit register information is available, making it possible to supplement the information in the SMF with data on bank-firm relations.

The CR information is extracted from the public Credit Register which is administered by the Bank of Italy; it pools information on the loans extended by Italian banks. Banks must report to the credit register the amounts granted and effectively utilized for all loans in excess of a minimum threshold which, at the relevant time for this paper, was 80 million lire (about $46,000). This information is reported separately for six types of loans (financial and commercial paper, foreign credit operations, credit lines, collateralized loans, medium- and long-term loans, and personal guarantees).

Finally, the BB data set retrieves information on the banks lending to each individual firm from a database of the population of Italian banks. This database contains both balance sheet variables and non-balance sheet bank characteristics, such as the number of branches, year of foundation, location of headquarters, number of employees, and bank ownership.

B. Explanatory Variables

Firm size: Logarithm of the number of firm employees in 1992. Source: SMF.

Firm leverage: Ratio of total sales in 1994 to beginning-of-period total bank debt (value of utilized bank loans). Source: SMF (sales) and CR (bank debt).

Share of defaulted loans recovered: Average share of defaulted loans that is recovered in the province where the firm is located. The figures come from a 1994 survey of loan recovery procedures carried out by the Bank of Italy on a sample of 269 banks (see Generale and Gobbi (1996)) for a description).

Nonperforming loans: Average share of nonperforming loans on loanable funds across relationship banks. The indicator is constructed as follows: for each bank in the population (about 850) and for each quarter between 1987 and 1994 the ratio between defaulted loans and loanable funds is computed; the ratio is then averaged across quarters to obtain a measure of the incidence of defaults for each bank. For each firm in the sample we then compute the average value of this indicator, using as weights the share of total bank loans contributed by each bank. Source: BB and CR.

Liquidity shocks: Using the BB database, for each bank in the population and each quarter from 1987 to 1995 the ratio between liquidity and loanable funds is computed. Liquidity is defined as *liq* = *free reserves* + 0.05 *required reserves* + *government bonds* + *net position in the interbank market*. We take five percent of the required reserves since this is the share that can be mobilized to face liquidity needs; for each bank i at time t the following variable is defined:

$$dliq(i,t) = [\log(liq(i,t)/mliq(i)]/se(i),$$

where $ml(i)$ is the mean across time of $\log(liq(i,t))$ and $se(i)$ is its standard error. This procedure accounts for bank fixed effects and for the possible presence of outliers. To eliminate the aggregate component of the liquidity shocks, we regress $dliq(i,t)$ on a set of time dummies, one for each quarter; the regression residual is taken as a measure of idiosyncratic liquidity shocks. For each bank, the standard error of this residual is then computed to obtain a measure of fragility. Finally, for each *firm*, we compute the average value of the standard error of the idiosyncratic liquidity shocks among the pool of its relationship banks, using as weights the share of loans extended.

Average bank size: The logarithm of the average size of each firm's relationship banks in terms of loanable funds, using as weights the shares of the total loan extended to the firm by each bank in the pool. Source: BB and CR.

Firm age: Computed as 1992 minus the year of foundation. Source: SMF.

Group membership: Dummy variable: equal to one if the firm belongs to a group, equal to zero otherwise. Source: SMF.

Industry comovement: The correlation among the sales of the firm and those of other firms in the same industry constructed by Guiso and Parigi (1999). The source of the data is a balanced panel of 5,054 Italian industrial firms from 1983 to 1995. To construct the indicator, the authors classify firms into 44 industrial sectors. For each sector, they regress the standardized rate of growth of firms sales (in deviation from its mean) on a set of year dummies, as measures of aggregate shocks. Sectors where aggregate shocks are more important (and thus where firm sales tend to co-vary closely) have a higher R^2 for this regression relative to sectors where idiosyncratic disturbances are dominant. These R^2 are then assigned to 1994 SMF firms on the basis of their industry code.

Share of first owner: Share of firm equity in the hands of largest shareholder. Source: SMF.

Indicators of R&D and adoption of innovations: The patents indicator is a dummy variable that is equal to one if the firm has either sold or purchased a patent; the R&D indicator is a dummy that is equal to one if the firm has invested in R&D in the three years covered by the SMF

(1992 to 1994); finally the three dummies for the adoption of innovations take the value one if during the period 1992 to 1994 the firm has innovated its products, production processes, or both.

REFERENCES

Asarnow, Elliot, and David Edwards, 1995, Measuring loss on defaulted bank loans: A 24-year study, *Journal of Commercial Lending* 77, 11–23.

Berger, Allen N., and Gregory F. Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of Business* 68, 351–381.

Berglöf, Erik, and Ernst-Ludwig Von Thadden, 1994, Short-term versus long-term interests: Capital structure with multiple investors, *Quarterly Journal of Economics* 109, 1055–1084.

Bolton, Patrick, and David S. Scharfstein, 1996, Optimal debt structure and the number of creditors, *Journal of Political Economy* 104, 1–25.

Buchinsky, Moshe, and Oved Yosha, 1995, Evaluating the probability of failure of a banking firm, Cowles Foundation Discussion paper no. 1108, Yale University.

Detragiache, Enrica, 1994, Public versus private borrowing: A theory with implications for bankruptcy reform, *Journal of Financial Intermediation* 3, 327–354.

Diamond, Douglas W., 1991a, Monitoring and reputation: The choice between bank loans and directly placed debt, *Journal of Political Economy* 99, 689–721.

Diamond, Douglas W., 1991b, Debt maturity structure, and liquidity risk, *Quarterly Journal of Economics* 106, 709–737.

Generale, Andrea, and Giorgio Gobbi, 1996, Loan recovery: cost, duration, and bank behavior, Temi di Discussione no. 265, Banca d'Italia.

Guiso, Luigi, and Giuseppe Parigi, 1999, Investment and demand uncertainty, *Quarterly Journal of Economics* 114, 185–227.

Harris, Milton, and Artur Raviv, 1991, Theory of capital structure, *Journal of Finance* 46, 297–355.

Houston, Joel, and Christopher James, 1996, Bank information monopolies and the mix of private and public debt claims, *Journal of Finance* 51, 1863–1889.

Lummer, Scott L., and John J. McConnell, 1989, Further evidence on the bank lending process and the capital market response to bank loan agreements, *Journal of Financial Economics* 25, 99–122.

Ongena, Steven, and David C. Smith, 1998, What determines the number of bank relationships? Cross-country evidence, Unpublished manuscript, Norwegian School of Management.

Petersen, Mitchell A., and Raghuram G. Rajan, 1994, The benefits of lending relationships: Evidence from small business data, *Journal of Finance* 49, 3–37.

Petersen, Mitchell A., and Raghuram G. Rajan, 1995, The effect of credit market competition on lending relationships, *Quarterly Journal of Economics* 110, 407–443.

Rajan, Raghuram G., 1992, Insiders and outsiders: The choice between informed and arm's-length debt, *Journal of Finance* 47, 1367–1400.

Sharpe, Steven, 1990, Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships, *Journal of Finance* 45, 1069–1087.

Shleifer, Andrei, and Robert W. Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance* 43, 1343–1366.

Von Thadden, Ernst-Ludwig, 1994, The commitment of finance, duplicated monitoring, and the investment horizon, Unpublished manuscript, Basel University.

Yosha, Oved, 1995, Information disclosure costs and the choice of financing source, *Journal of Financial Intermediation* 4, 3–20.

THE JOURNAL OF FINANCE • VOL. LX, NO. 6 • DECEMBER 2005

On the Benefits of Concurrent Lending and Underwriting

STEVEN DRUCKER and MANJU PURI*

ABSTRACT

This paper examines whether there are efficiencies that benefit issuers and underwriters when a financial intermediary concurrently lends to an issuer while also underwriting its public securities offering. We find issuers, particularly noninvestment-grade issuers for whom informational economies of scope are likely to be large, benefit through lower underwriter fees and discounted loan yield spreads. Underwriters, both commercial banks as well as investment banks, engage in concurrent lending and provide price discounts, albeit in different ways. We find concurrent lending helps underwriters build relationships, increasing the probability of receiving current and future business.

FOR MANY YEARS, THE 1933 GLASS-STEAGALL ACT prevented commercial banks from underwriting corporate bonds and equities. Due to the relaxation and recent repeal of the Act, many commercial banks have acquired investment banks or developed investment banking capabilities internally to create universal banks that can offer an array of financial services.

The entry of commercial banks into underwriting markets has increased the potential for financial institutions to offer both lending and underwriting services. In particular, it has become increasingly common for financial intermediaries to provide loans to a firm while also underwriting the firm's public securities. In fact, concurrent lending and underwriting has increased substantially over time—in 1994, only 1% of seasoned equity issuers received a loan from their underwriter at around the time of issuance, but by 2001, over 20% of all deals were concurrent. The movement toward concurrent lending and underwriting raises a host of interesting questions. First, why are deals concurrent?

*Drucker is from the Graduate School of Business, Columbia University. Puri is from the Fuqua School of Business, Duke University and NBER. Puri is grateful to the Sloan Foundation for partial funding of this research. This paper was circulated earlier under the title "Tying knots: Lending to win equity underwriting business." We thank James Booth, Mark Flannery, Jean Helwege, Jianping Qi, Tim Loughran, Rob Stambaugh (the editor), Steven Ongena, Andy Winton, an anonymous referee, and seminar participants at the Board of Governors, Washington D.C., the Federal Reserve Bank of San Francisco, Indiana University, INSEAD, London Business School, Stanford Graduate School of Business, the 2003 Financial Management Association International and Annual Meetings, the 2003 Federal Deposit Insurance Corporation Conference on Finance and Banking, the 2004 American Finance Association Meeting, the 2004 International Industrial Organization Conference, the 2004 Federal Reserve Bank of Chicago Conference on Bank Structure & Competition, and the 2004 Financial Intermediation Research Society Conference on Banking, Insurance, and Intermediation.

Are there efficiencies resulting from offering lending and underwriting services at the same time? Concurrent deals might allow for potential efficiency gains due to informational economies of scope that can result from the bank jointly delivering services and using the same client-specific information for multiple purposes (see, e.g., Benston (1990), Saunders and Walter (1994)). Therefore, concurrent lending and underwriting might be useful in cases in which there are large potential economies of scope from combining lending and underwriting. This would suggest that certain kinds of deals are concurrent but not others. Second, who benefits from concurrent lending and underwriting? Lower costs could arise due to informational economies of scope, and issuers could benefit if the bank passes along these savings. For the underwriter, providing concurrent lending and underwriting services might help build relationships that improve the probability of securing current and future business from the firm. Third, do the benefits from concurrent lending and underwriting vary by the type of underwriter involved in the transaction? It is possible that commercial banks are able to generate larger economies of scope than investment banks due to their well-established lending businesses, and therefore, there may exist differences in concurrent deals that are underwritten by investment banks as opposed to commercial banks.

In this paper, we address these issues empirically by studying instances in which underwriters concurrently lend to firms and underwrite these firms' seasoned equity offerings (SEOs). To tackle these questions, we use a unique data set that is carefully assembled from multiple databases and augmented by hand-collected data. We gather data on seasoned equity issuers, including each firm's credit rating, stock returns, issuance history, and lending history. We identify prior underwriting and lending relationships between each issuer and potential underwriter, as well as each underwriter's ranking, level of analyst coverage, and quality of analyst coverage. Further, we collect data on underwriter fees, loan pricing, and lending terms.

We find that there is a distinct profile of issuers that are involved in concurrent deals. In the majority of concurrent deals, the firms are highly leveraged and noninvestment-grade rated. One explanation for this is that for lower rated and highly leveraged firms, there are larger potential efficiency gains that arise due to informational economies of scope from combining lending and underwriting. Therefore, concurrently offering lending and underwriting for these issuers could produce substantial benefits. To study whether issuers actually benefit from the concurrent offering of services, we examine the impact on issuers' financing costs. Our results suggest that concurrent lending and underwriting lowers issuers' financing costs in two ways: (i) a reduced underwriter fee for the equity offering and (ii) discounted yield spreads of concurrent loans as compared with "matched" nonconcurrent loans. Interestingly, we find that the cost reductions are more pronounced among issuers that are noninvestment-grade rated, for whom the expected informational economies of scope are relatively large. Further, concurrent deals where the firm and underwriter have a prior lending relationship produce larger underwriter fee reductions, which is again consistent with the existence of scope economies between lending and underwriting. In fact, prior lending relationships, in general, are associated with significantly lower underwriter fees.

To ensure that matching biases are not driving the yield spread discount, we use the econometric techniques developed by Heckman, Ichimura, and Todd (1997, 1998). These econometric methods effectively take into account the fact that the characteristics of concurrent loans may differ significantly from nonconcurrent loans and ensure that such observed differences are not driving the results. Using a variety of matching models, we confirm that concurrent loans are significantly cheaper than comparable loans.

To examine whether underwriters benefit from offering loans at the same time as an equity issuance, we look at the impact of concurrent deals on the underwriter's relationship with the firm. In particular, we investigate whether the same bank is selected for current and future equity underwriting mandates. We find that lending at the time of a seasoned equity issuance significantly increases the probability of securing current equity underwriting business. Concurrent lending is an important factor in the selection of an underwriter both when firms and underwriters have not interacted through previous lending transactions and also when they do have a prior lending relationship. We also find that issuers that have received a concurrent loan during a previous SEO return to the equity market more frequently than nonconcurrent issuers and that issuers who were involved in a prior concurrent deal with an investment bank underwriter are more likely to keep the same underwriter. The significant effects of concurrent lending on an underwriter's ability to generate both current and future underwriting business hold even after controlling for issuer characteristics and other factors that are likely to affect underwriter selection, such as underwriter reputation, prior relationships, and the level and quality of analyst coverage provided by underwriters. Further, our estimations show that prior lending relationships are important factors in determining underwriter selection in both current and future equity offerings. These results are consistent with concurrent loans, and more generally, the use of lending to help build relationships that increase an underwriter's expected revenues.

Lastly, we examine whether the benefits from concurrent deals vary by the type of underwriter involved in the transaction. Interestingly, while commercial banks are well positioned to offer lending and underwriting services concurrently due to their existing lending businesses, we discover that investment banks underwrite a significant portion of concurrent deals. This suggests that investment banks have developed the organizational infrastructure to lend and is consistent with there being potential gains from a single entity offering both lending and underwriting services.[1] Our results thus indicate that commercial banks and investment banks both compete for concurrent

[1] For example, Morgan Stanley participated in a $6.5 billion bank loan for Lucent Technologies and was subsequently awarded the role of underwriter on Lucent's spinoff of Agere Technologies (see Smith (2001a)). Moreover, investment banks are increasing their lending capacity, with Merrill Lynch, Lehman Brothers, and Morgan Stanley forming bank subsidiaries (see Smith (2001b)).

deals; however, they seem to compete through different components of the concurrent deals—commercial banks are more likely to offer discounted yield spreads on concurrent loans, while investment banks are more likely to discount the underwriter spread for the SEO. This is consistent with each type of underwriter competing more aggressively in its area of expertise and in the area in which it is more likely to generate future business: investment banks discount underwriter spreads and receive more future underwriting business; commercial banks discount loan yield spreads, which is consistent with establishing a lending relationship that helps generate future banking business.

This paper adds to the growing literature on how underwriters and issuers associate with each other. An important question is what determines the pairing of firms and underwriters for current as well as future deals? Studies suggest that underwriter reputation is an important determinant of the choice of underwriter (Booth and Smith (1986), Carter and Manaster (1990)), and high-quality issuers are more likely to associate with high-quality underwriters (Fernando, Gatchev, and Spindt (2005)). Underwriter capability in terms of all-star analyst coverage has been found to be important in affecting investment banking deal flow (Clarke et al. (2003), Corwin and Schultz (2005)) and for switching from one underwriter to another (Krigman, Shaw, and Womack (2001)), though there is little evidence to suggest that aggressive analyst recommendations increase the bank's probability of winning an underwriting mandate (Ljungqvist, Marston, and Wilhelm (2004)). In this paper, we find that concurrent lending as well as prior lending by the underwriter to the firm significantly affects firm–underwriter pairings and the pricing of underwriting services. Lending activities are important not just for current firm–underwriter association but also for future transactions and help create durable relationships that can benefit the issuer through lower financing costs.[2] Our findings also underscore that firm–underwriter pairings can differ by underwriter type and not simply by underwriter reputation and analyst coverage, as we find important differences between commercial bank and investment bank underwriters.

This paper also contributes to the literature on universal banking and the implications of allowing banks to underwrite securities. Regulators have recently raised questions on the firm-level and competitive effects of the relaxation and repeal of the Glass-Steagall Act (see, e.g., Berger, Demsetz, and Strahan (1999), Santomero and Eckles (2000)). Allowing banks to both lend and underwrite raises many concerns, including the potential for banks to engage in tying practices, where financial institutions alter the pricing or provision of credit based on a firm's decision to use the bank's investment banking services.[3] Since it is illegal for commercial banks to tie lending to underwriting services, explicit

[2] See also Ljungqvist et al. (2004) for additional sources of durability in bank–issuer relationships.

[3] U.S. House Representative Dingell highlights some regulatory concerns in a letter to Chairman Greenspan and Comptroller Hawke (see Dingell (2002)).

agreements are rarely found.[4] However, to the extent that concurrent lending and underwriting proxies for implicit agreements between underwriters and firms, our results suggest that firms actually benefit from using the same underwriter to arrange both of the concurrent transactions. Regarding the entry of commercial banks into underwriting, the theoretical literature has examined the potential for commercial banks and investment banks to coexist, as well as the implications of such a scenario (see, e.g., Boot and Thakor (1997), Kanatas and Qi (1998, 2003), Puri (1999), Rajan (2002), Stefanadis (2004)). However, the possibility that investment banks might respond by expanding into lending activities has generally not received much attention. Our results bring to light some similarities and differences in the ways in which investment banks and commercial banks compete for underwriting business. We also add to the evidence on implications of combining lending with underwriting. Much of the empirical literature that examines when banks lend and underwrite investigates the effect of bank lending, and the private information contained therein, on the banks' underwriting of public securities.[5] These effects are ascertained through the pricing of underwritten securities (see, e.g., Puri (1996), Gande et al. (1997), Yasuda (2005), Benzoni and Schenone (2004)) or through long-run performance (see, e.g., Ang and Richardson (1994), Kroszner and Rajan (1994), Puri (1994)). An important but unexplored issue is the reverse question—how do potential underwriting opportunities affect banks' lending, and how does this in turn affect the financing cost of the issuing firm? This paper provides a first step in addressing this question.

The remainder of the paper is organized as follows. Section I describes the data and our sample selection process. We present the major empirical findings in Section II. Section III concludes.

I. Data and Sample Selection

We attempt to capture instances in which a financial institution lends to a firm and concurrently underwrites its public security issuance. The definition that we adopt is if the firm receives a loan from the underwriter of the SEO

[4] Section 106 of the Bank Holding Company Act Amendments of 1970 prohibits a bank from explicitly extending credit or varying the terms of credit on the condition that a customer purchase another product or service from the bank or its affiliates. However, the Federal Reserve recently stated that the laws "do not prohibit a bank from granting credit or providing any other product to a customer based solely on a desire or a hope (but not a requirement) that the customer will obtain additional products from the bank or its affiliates in the future." Also, clients are free to use "their own bargaining power" to seek a bundle of banking services. For more information, see Michaels and Silverman (2003).

[5] In related literature, James (1987), Lummer and McConnell (1989), Best and Zhang (1993), and Billett, Flannery, and Garfinkel (1995), among others, find that new loans, loan renewals, and lender identity carry (positive) private information to the outside equity market about a borrowing firm's financial condition. See James and Smith (2000) for a comprehensive review of the past and recent research on the special nature of bank loan financing. This literature examines the effect of bank lending absent an underwriting role for the bank.

Table I

Concurrent Deals, by Year

This table presents the percentage of seasoned equity offerings (SEOs) that are concurrent deals. A concurrent deal is any SEO in which the underwriter provides a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO.

Year	1994	1995	1996	1997	1998	1999	2000	2001[a]
Number of SEOs	363	493	596	515	340	389	375	86
Number of concurrent deals	5	5	19	48	37	52	27	18
% Concurrent deals	1.38	1.01	3.19	9.32	10.88	13.37	7.20	20.93

[a]Through May 31.

between 6 months prior to and 6 months after the SEO, we classify the loan as a "concurrent loan" and the SEO as a "concurrent deal." As a robustness check to this definition, we also run our estimations where we define concurrent loans to be those loans that were originated between 3 months prior to and 3 months after the SEO. This sample produces qualitatively similar results.

We select our sample period based on the following factors. First, we hope to capture an active period of concurrent lending and underwriting. Table I shows that concurrent deals were nearly nonexistent before 1996 and with the exception of the year 2000, the proportion of concurrent deals increases each year. The decline in concurrent deals in the year 2000 may be due to a noticeable decline in telecom and cable SEOs, which account for around one third of all concurrent deals, and a very high proportion of technology offerings, which account for only a small percentage of concurrent deals. Second, since we will be examining whether the issuers proceed with a subsequent SEO, we must provide enough time to capture the decisions of end-of-sample issuers. Based on these considerations, we define our sample period as January 1, 1996 through May 31, 2001.

We construct a unique database using eight different data sources and hand-collected data. All variables are defined in detail in Appendix A. Data on SEOs comes from Thomson Financial's SDC *Platinum* United States New Issues database, from which we download underwritten, seasoned, U.S. common stock issues. Since we wish to study industrial firms, we remove financial firms (companies with a one-digit SIC code of six). The sample consists of 2,301 issues. We hand match, by issuer name, each of the 2,301 issuers to the Loan Pricing Corporation's (LPC) *DealScan* database to identify whether the firm received a concurrent loan from their underwriter, and in doing so, we identify whether the SEO is a concurrent deal.[6] There are 201 concurrent deals in the sample and 2,100 nonconcurrent deals.

[6] LPC *DealScan* collects its loan data from SEC filings, and it receives data from large loan syndicators and from a staff of reporters. As such, *DealScan* is well suited to studying the borrowing activity of companies with public equity and debt. Since all of the companies in our sample have public equity, we should observe the vast majority of their lending activity. *Dealscan* has been used in previous studies for many purposes, including examining the effect of lending on bond yield spreads (see, e.g., Gande et al. (1997)) and bank effects in lending rates (Hubbard Kuttner, and Palia (2002)).

We classify each underwriter as an "investment bank" or a "commercial bank" based on the status of the parent/holding company of the underwriter at the time of the issue.[7] Due to the many mergers and acquisitions in the financial sector, we use the mergers and acquisitions database from SDC *Platinum* to aid in classification. For example, NationsBank acquired Montgomery Securities on October 1, 1997. Montgomery Securities is classified as an investment bank prior to October 1, 1997, but after October 1, 1997, we classify it as a commercial bank. Commercial banks underwrote 91 concurrent SEOs and 591 nonconcurrent SEOs, while investment banks underwrote the remaining 110 concurrent SEOs and 1,509 nonconcurrent SEOs.

As we study how concurrent lending and underwriting affect the pricing of bank services and the ability of the underwriter to generate equity underwriting business, we need to control for factors that may alter fees, pricing, or the likelihood that an issuer selects an underwriter. Prior underwriting relationships are likely to be important in both the selection of a bank and the pricing of banking services (see, e.g., Baker (1990), James (1992), Crane and Eccles (1993), Ljungqvist et al. (2004)). Furthermore, it is possible that prior lending relationships could also influence underwriter selection and the pricing of services. In particular, if there are economies of scope in lending and underwriting, then a prior lending relationship may result in a reduced underwriter fee or other pricing differences. When identifying prior lending and underwriting relationships, we account for mergers between potential underwriters. For example, Fleet Bank merged with BankBoston/Robertson Stephens on October 1, 1999. When tracking relationships, we assume that Fleet Bank acquired all of BankBoston's and Robertson Stephens' prior lending and underwriting relationships. From SDC *Platinum*, we identify 90 concurrent issuers and 830 nonconcurrent issuers that use an underwriter that had underwritten a prior equity offering. From *DealScan*, we identify 83 concurrent issuers and 103 nonconcurrent issuers that have a prior lending relationship with the selected underwriter.

Previous research indicates that we need to incorporate the reputation of the underwriter, the level of analyst coverage, and the quality of analyst coverage into our models because these factors are likely to affect the firm's decision to select an underwriter or to switch underwriters in the future. We capture the influence of reputation through the underwriter's market share. For each year, we compute each underwriter's SEO market share by adding the principal amounts of all SEOs in which the bank was the underwriter and dividing this total by the principal amounts of all SEOs during the year. If a merger between underwriters occurred during the year, we use the combined market share of the underwriters. We rank the underwriters on a yearly basis, based

[7] We do not separate commercial banks that internally developed investment banking capabilities from those that acquired investment banks because almost all of the commercial banks developed underwriting operations by acquiring investment banks. Chaplinsky and Erwin (2001) note that for commercial banks who developed underwriting capabilities internally, only JP Morgan acquired market share in equity underwriting that is above 0.02% during the post-1996 period.

on the market share in the previous year.[8] For example, Goldman Sachs had the highest market share in 1995, so in our models, issuers who have an SEO in 1996 consider Goldman Sachs to be the top-ranked underwriter.

We measure the level of equity analyst coverage by using the I/B/E/S Detail History, which contains over 12 years of forecast changes and encompasses earnings estimates from more than 200 brokerage houses and 2,000 individual analysts. We match any estimate of earnings per share from any analyst in the I/B/E/S database to each of the 2,301 firms in our sample. If the underwriter provided an earnings recommendation within 1 year prior to the SEO date, then the underwriter provided "coverage." To capture the quality of analyst coverage, we use Institutional Investor magazine's All-America Research Team, which is published yearly and lists the top-three analysts in each sector. Since the report is published toward the end of each year, the inclusion of an analyst in the publication will most likely have its greatest impact on underwriter choice for issues that occur in the following year. As a result, for our purposes here, we say that the analyst (and corresponding underwriter) provided "all-star coverage" for a firm if the analyst is included in the All-America Research Team for the year prior to the equity issuance and provided an earnings recommendation within 1 year prior to the SEO date.

Since it is necessary to control for financial characteristics and risk factors, we obtain financial data for each firm from the Compustat Industrial Quarterly database from Standard and Poor's. The financial data used in this study correspond to the quarter in the year of the SEO issue date. The incorporation date for each firm is hand collected from Moody's/Mergent's Industrial and Transportation Manuals and Standard & Poor's Corporation Records. From the Center for Research in Security Prices daily stock database, we download daily return, price, and outstanding share data to compute the equity volatility and market capitalization for each firm.

For each of the 201 concurrent deals, we gather the associated lending facilities from LPC *DealScan*. There are 358 concurrent lending facilities. The sample of concurrent lending facilities consists of 116 notes, 111 revolving lines of credit, 99 term loans, seventeen 364-day facilities, 13 bridge loans, and 2 other types of facility.

To examine differences between concurrent loans and nonconcurrent loans, we create two separate samples. In the hand-matching sample, for each of the concurrent loan facilities, we create a control group of nonconcurrent loans that were originated at around the same time as the concurrent loan, with firms that belong to the same industry and have the same credit rating. We use all loans in *DealScan* that occur between 6 months prior to and 6 months after the term facility active date of the concurrent loan.[9] We keep only those nonconcurrent

[8] A simultaneity problem would arise if we used the market shares from the current year to rank the underwriters because when an issuer selects an underwriter in the current year, the decision simultaneously increases the underwriter's market share.

[9] We also use a sample of loans that occurs between 3 months prior to and 3 months after the SEO date. Results using this sample are similar and are not reported.

loans that have the same two-digit SIC code and credit rating as the corresponding concurrent loan. We remove any loan that is missing information for the all-in spread drawn and/or the length of the loan.[10] All bridge loans and loans with an issuer that is not rated are removed. This sample has 107 concurrent loans that can be matched to a similar nonconcurrent loan and comprises 56 revolving lines of credit, 40 term loans, ten 364-day facilities, and 1 other type of facility.

To construct the econometric-matching sample, we download all lending facilities in *DealScan* that occur between January 1, 1996 and May 31, 2001. We remove any facility that is missing information for the all-in spread drawn and/or the length of the facility, and we remove any facility such that the borrower is a financial firm (companies with a one-digit SIC code of six). As before, all bridge loans and loans to not-rated borrowers are excluded. This sample consists of 166 concurrent loans that can be matched to a sample of 6,919 nonconcurrent loans. Seventy-four revolving lines of credit, 77 term loans, fourteen 364-day facilities, and 1 other type of facility form the sample of 166 concurrent loans. Seventy-nine of the 166 concurrent loans are from commercial bank underwriters, while investment bank underwriters provide the remaining 87 concurrent loans.

In addition, we classify 340 lending facilities as "other issuance period loans," which are loans to an issuer of an SEO that are originated between 6 months prior to and 6 months after the SEO, where the lender could have been selected to underwrite the SEO but is not provided with underwriting responsibilities. Of the 6,919 nonconcurrent loans in the econometric-matching sample, 145 lending facilities are designated as other issuance period loans.

II. Methodology and Results

As shown in Table I, concurrent deals increased over time from about 1% in 1994 to over 20% in 2001. However, before 1996, while concurrent deals were nearly nonexistent, many issuers received loans from another bank at about the same time as the issuance of public securities.[11] Over time, issuers have shifted from using a commercial bank for lending and an investment bank for equity underwriting to employing a single entity for both of these transactions.

Table II reports summary statistics for the concurrent and nonconcurrent SEO samples. Concurrent issuers are highly leveraged, with debt-to-equity ratios that are, on average, five times higher than nonconcurrent issuers. Furthermore, concurrent borrowers have low credit ratings, with 71% of investment bank concurrent deals and 60% of commercial bank concurrent deals for junk-rated issuers, and another 12% of investment bank deals and 27% of

[10] The all-in spread drawn is the rate the borrower pays to the lender each year for each dollar drawn off the credit line (inclusive of fees), quoted in basis points over LIBOR.

[11] In 1994, over 30% of SEO issuers received a loan from some bank within a period of 6 months before and 6 months after the issuance, even though only 1.4% of these loans came from the underwriter of the issuance.

Table II

Univariate Tests for Differences in the Sample of SEOs between January 1996 and May 2001

This table tests for differences between concurrent deals and nonconcurrent deals and for differences between investment bank concurrent deals and commercial bank concurrent deals. Panels A and C use a difference in means t-test and Wilcoxon rank test. A concurrent deal is any seasoned equity offering (SEO) in which the underwriter provides a loan to the issuer between six months prior to the SEO and six months after the SEO. The underwriter is an IB (CB) if the parent or holding company of the underwriter is an investment bank (commercial bank) at the time of the SEO. The variables are defined as follows: USPREAD is the underwriter spread, expressed as a percentage of the principal amount; LNSIZE is the logarithm of the SEO principal amount, expressed in millions of dollars; DE-LTDEBT is the long-term debt to common equity ratio in the quarter of the SEO; AGE is the firm's age, measured in years; PRIORLEND is one if a loan between the underwriter and the issuer was originated at any time before six months prior to the SEO; PRIORUND is one if the underwriter had been the underwriter on any prior equity offering by the issuer; COVERAGE is one if the underwriter had provided an earnings per share estimate for the firm within the year prior to the SEO; and, ALLSTAR is one if COVERAGE is 1 and the analyst was ranked as an all-star by Institutional Investor magazine for the year prior to the SEO. A firm has an issuer rating of IGRADE if it is rated AAA, AA, A, or BBB by Standard & Poor's in the quarter of the SEO. A firm has an issuer rating of JUNK if it is rated BB, B, CCC, CC, or C by Standard & Poor's in the quarter of the SEO. All variables are explained in detail in Appendix A. ***, **, * indicates significantly different than zero at the 1%, 5%, and 10% level, respectively.

Panel A: Concurrent vs. Nonconcurrent Deals—Issuer and Issuance Variables

Variable	Concurrent Deal Mean	Nonconcurrent Deal Mean	*T*-Ratio	Wilcoxon Test *p*-Value
USPREAD	4.33	5.11	−8.63***	0.0000***
LNSIZE	5.09	4.28	9.94***	0.0000***
DE-LTDEBT	2.57	0.55	2.96***	0.0000***
AGE	21.78	17.87	2.12**	0.1845

Panel B: Concurrent vs. Nonconcurrent Deals—Relationship Variables

Variable	Percent of Concurrent Deals	Percent of Nonconcurrent Deals
CB	45.3	28.1
IB	54.7	71.9
PRIORLEND	41.3	4.9
PRIORUND	44.8	39.5
COVERAGE	77.1	63.0
ALLSTAR	21.4	12.9

Panel C: IB vs. CB Concurrent Deals—Issuer and Issuance Variables

Variable	IB Concurrent Deal Mean	CB Concurrent Deal Mean	*T*-Ratio	Wilcoxon Test *p*-Value
USPREAD	4.25	4.43	0.98	0.2792
LNSIZE	5.28	4.92	2.24**	0.0110**
DE-LTDEBT	2.83	2.31	0.39	0.4189
AGE	20.50	23.35	0.79	0.1148

(*continued*)

Table II—*Continued*

Panel D: IB vs. CB Concurrent Deals—Relationship Variables		
Variable	Percent of IB Concurrent Deals	Percent of CB Concurrent Deals
PRIORLEND	36.4	47.3
PRIORUND	48.2	40.7
COVERAGE	78.2	75.8
ALLSTAR	23.6	18.7

Panel E: IB vs. CB Concurrent Deals—Issuer Rating		
Variable	Percent of IB Concurrent Deals	Percent of CB Concurrent Deals
IGRADE	17.27	13.19
JUNK	70.91	60.44

commercial bank deals involving issuers that are not rated. Since duplication of information is particularly costly for risky firms because they are subject to extensive due diligence in both lending and underwriting, concurrent lending and underwriting can be extremely beneficial for these issuers because a single bank can use the collected information for both transactions. In addition, for lower rated and highly leveraged firms, debt has similar characteristics to equity. As a result, information gathered in the lending process will be relevant to the equity issuance that may enhance the certification ability of the underwriter. Therefore, economies of scope are likely to be high for these firms, and concurrent lending and underwriting may be an efficient response to the ability of the banks to use information across product lines.

Commercial banks are underwriters on 45% of concurrent deals and investment banks underwrite the remaining 55% of concurrent deals. Also, commercial banks and investment banks provide concurrent loans to similar clients. These are interesting facts, which suggest that investment banks have now developed the organizational structure to lend. This expansion into lending by investment banks is consistent with there being potential gains from a single entity offering both lending and underwriting services.

A. Equity Underwriter Spreads

We wish to determine whether concurrent lending and underwriting lowers issuers' financing costs. One possibility is that the firm pays a lower fee to the bank for underwriting its equity offering. An underwriter could charge a lower fee in a concurrent deal because the bank may face lower underwriting costs due to informational economies of scope that arise from the joint delivery of services and the reusability of information gathered during the lending process. We examine differences between concurrent and nonconcurrent underwriting fees by analyzing the underwriter spread, which is the compensation paid to the underwriter for selling the firm's security issue, expressed as a percent of the capital raised. Consistent with the existence of scope economies, the univariate descriptive statistics in Table II, Panel A, indicate that the average underwriter spread of concurrent SEOs is 78 basis points lower than the mean underwriter spread of nonconcurrent SEOs, a difference that is significant at the 1% level.

A.1. U-Shaped Underwriter Spreads

The initial evidence indicates that concurrent issuers receive lower underwriter spreads. We wish to see if this result withstands a multivariate specification. Following Altinkilic and Hansen (2000), we estimate a model of the underwriter spread that can be a U-shaped function of the amount of new capital raised. Theoretically, a U-shaped function could arise because fixed costs cause scale economies initially but as issue size increases, diseconomies of scale arise in the spread due to rising placement costs. Altinkilic and Hansen find strong evidence of U-shaped curves in a sample of 1,325 SEOs from 1990 through 1997.

As a model for the underwriter spread, we use Altinkilic and Hansen's (2000) expanded spread model in which the underwriter spread is the sum of a fixed cost and a variable cost component. In order to generate U-shaped spreads, the variable cost component must be allowed to rise over a relevant range of proceeds. This condition is satisfied by dividing the SEO principal amount by the firm's equity market capitalization, which effectively holds the firm size fixed as the size of the offering expands, thereby allowing variable costs of underwriting to increase at an increasing rate. We control for the volatility of equity returns because higher volatility can cause more uncertainty, which may be reflected in a higher underwriter spread. The model captures any variation in underwriter costs that is due to the volume of issuance in the seasoned equity market.

We extend the model to include variables to capture concurrent lending and prior relationships. Since an existing lending relationship can lower setup costs and provide the bank with access to additional information, concurrent deals involving prior lenders may be less costly. To capture this potential effect, we control for interactions between prior lending and concurrent lending. A negative coefficient on the concurrent lending variables would be consistent with the existence of scope economies. We estimate two variations of the expanded spread model—in the first model, we do not consider differences between investment banks and commercial banks; we relax this restriction in the second model. Further, we examine differences between noninvestment-grade and investment-grade issuers. Since economies of scope are likely to be high for noninvestment-grade firms, we expect discounts to be concentrated among these deals.

A.2. Results

Results of ordinary least squares regressions are presented in Table III. We find support for U-shaped spreads. As more capital is raised, the variable cost increases. As expected, higher stock return volatility increases the variable

Table III

Underwriter Spread Regressions

This table provides ordinary least squares estimates of a model of the underwriter spread that can be a U-shaped function of the amount of new capital raised. The model is based on Altinkilic and Hansen's (2000) expanded spread model. The dependent variable is USPREAD, the underwriter spread, expressed as a percentage of the principal amount. The independent variables are as follows: CONCLOAN is a dummy variable that equals 1 if the underwriter provides a loan to the issuer between 6 months prior to the seasoned equity offering (SEO) and 6 months after the SEO *and* the underwriter has never provided a loan to the issuer in the past; CONCPLEND is a dummy variable that equals 1 if the underwriter provides a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the underwriter provided a loan to the issuer prior to 6 months before the SEO; PRIORLEND is a dummy variable that equals 1 if a loan between the underwriter and the issuer was originated at any time prior to 6 months before the SEO *and* the underwriter does not provide a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO; PRIORUND is 1 if the underwriter had been the underwriter on any prior equity offering by the issuer; IB (CB) is 1 if the parent / holding company of the underwriter is an investment bank (commercial bank); 1/SEOSIZE is the inverse of the principal amount of the equity offering, measured in millions of dollars; SEOSIZE/MKTCAP is the principal amount of the offering divided by the market capitalization of the issuer at the date of the SEO; VOL is the daily standard deviation of the issuer's common stock rate of return; MKTACT is the dollar volume of issuance in the U.S. SEO market for the 3 months prior to each offering; and, SICx are industry dummy variables, which are one if the firm has the corresponding one-digit SIC. All variables are explained in detail in Appendix A. In columns (1) and (4), we estimate the models using the full sample of issues. In columns (2) and (5), the sample is restricted to SEOs by noninvestment-grade issuers. Noninvestment-grade issuers are either not rated or have a Standard & Poor's long-term debt rating of BB, B, CCC, or CC in the quarter of the SEO. In columns (3) and (6), the sample is restricted to SEOs by investment-grade issuers. Investment-grade issuers have a Standard & Poor's long-term debt rating of AAA, AA, A, or BBB in the quarter of the SEO. Coefficients for the industry variables (SICx) are not reported. *T*-ratios are in parentheses. ***, **, * indicates significantly different than zero at the 1%, 5%, and 10% level, respectively.

	Full Sample (1)	Noninvestment-Grade (2)	Investment-Grade (3)	Full Sample (4)	Noninvestment-Grade (5)	Investment-Grade (6)
Intercept	4.247***	4.599***	3.439***	4.231***	4.565***	3.185***
	(33.12)	(35.74)	(7.38)	(31.57)	(34.28)	(5.64)
CONCLOAN	−0.182*	−0.179*	−0.034			
	(−1.74)	(−1.76)	(−0.11)			
CONCPLEND	−0.360**	−0.329**	−0.474			
	(−2.31)	(−2.26)	(−1.02)			
PRIORLEND	−0.360***	−0.358***	−0.069			
	(−3.04)	(−3.17)	(−0.23)			

(continued)

Table III—*Continued*

	Full Sample (1)	Noninvestment-Grade (2)	Investment-Grade (3)	Full Sample (4)	Noninvestment-Grade (5)	Investment-Grade (6)
PRIORUND	−0.217***	−0.263***	−0.028			
	(−4.19)	(−5.19)	(−0.17)			
IB				0.021	0.043	0.238
				(0.29)	(0.64)	(0.67)
IB × CONCLOAN				−0.263**	−0.343**	0.303
				(−2.00)	(−2.40)	(0.92)
CB × CONCLOAN				−0.070	0.022	−0.760
				(−0.43)	(0.17)	(−1.61)
IB × CONCPLEND				−0.440**	−0.413**	−1.382
				(−2.20)	(−2.44)	(−1.55)
CB × CONCPLEND				−0.321	−0.283	0.003
				(−1.43)	(−1.27)	(0.00)
IB × PRIORLEND				−0.324**	−0.328***	0.046
				(−2.49)	(−2.67)	(0.15)
CB × PRIORLEND				−0.454*	−0.427*	−0.441
				(−1.81)	(−1.84)	(−0.48)
IB × PRIORUND				−0.248***	−0.299***	−0.122
				(−4.39)	(−5.29)	(−0.72)
CB × PRIORUND				−0.135	−0.173**	0.178
				(−1.45)	(−2.02)	(0.42)
1/SEOSIZE	17.270***	15.377***	24.680***	17.259***	15.328***	24.783***
	(6.04)	(5.99)	(2.91)	(5.98)	(5.92)	(2.74)
SEOSIZE/MKTCAP	0.242	0.225	−0.049	0.241	0.223	−0.052
	(1.43)	(1.13)	(−0.37)	(1.42)	(1.12)	(−0.39)
VOL	12.274***	7.570***	17.273	12.226***	7.532***	18.625*
	(10.26)	(6.73)	(1.57)	(9.96)	(6.55)	(1.73)
MKTACT	−7.581**	−4.071	−2.042	−7.652**	−3.957	−2.173
	(−2.34)	(−1.44)	(−1.38)	(−2.36)	(−1.42)	(−1.49)
R^2	0.4029	0.4003	0.1644	0.4040	0.4026	0.2048

spread and there is a large fixed cost component to underwriter spreads. In the first column of Table III, we present the results of the model in which we do not consider differences in the fees charged by investment banks and commercial banks. The coefficients on the concurrent lending and the prior lending variables are all negative and significant. A concurrent loan without a prior lending relationship provides an 18-basis point reduction in the underwriter spread, which is significant at the 10% level. A prior lending relationship, both with and without a concurrent loan, translates into a 36-basis point reduction in the underwriter spread. On a $200 million equity offering, an 18-basis point reduction in the underwriter fee provides a cost savings of $360,000 to the issuer, while a 36-basis point decrease saves the issuer $720,000. These results are consistent with the existence of economies of scope.

As previously argued, economies of scope between lending and underwriting are likely to be pronounced when the issuer is junk rated or not rated. We restrict the sample of SEOs to include only junk-rated and not-rated issuers and display the results of the model in the second column of Table III. Consistent with the existence of informational economies of scope, we find that among these issuers, significant underwriter spread discounts are provided when the issuer receives a concurrent loan or has a prior lending relationship with the underwriter. In the third column of Table III, we present the results of the model in which we restrict the sample to include only investment-grade issuers. Among investment-grade issuers, for whom private information is likely to be less important, we do not find significant underwriter spread discounts. These results highlight that the underwriter spread discounts are driven by deals in which, ex ante, concurrent lending and underwriting is likely to be efficient.

The results in the fourth and fifth columns of Table III show that investment banks account for most of the concurrent lending and underwriting discount. For concurrent issuers, investment banks provide a discount of 26 basis points if no prior lending relationship exists and 44 basis points if there is a prior lending relationship, both significant at the 5% level. On a $200 million equity offering with an investment bank, on average the issuer saves $520,000–$880,000. For commercial bank underwritten issues, the coefficients for concurrent deals are negative but insignificant. It is interesting to note that both investment banks and commercial banks provide significant discounts in the underwriter spread to firms that do not receive a concurrent loan but with which a prior lending relationship is in place, which further supports the existence of informational economies of scope between lending and equity underwriting.

Overall, we find that concurrent deals have lower underwriter spreads than nonconcurrent deals and that concurrent deals such that there was a prior lending relationship in place receive a larger discount. Importantly, we find that the discounts are driven by deals that involve junk-rated and not-rated issuers, for whom economies of scope between lending and underwriting are likely to be large. Consequently, the results are consistent with the view that concurrent deals are an efficient response to the ability of banks to use information across product lines. We find additional support for the existence of economies of scope between lending and equity underwriting, as a prior lending relationship translates into an underwriter spread discount. Further, we find that most of the underwriter spread discounts can be attributed to investment bank underwriters.

B. The Pricing of Concurrent Loans

We now study the pricing of concurrent loans to address two issues. First, we wish to determine whether there is additional evidence that concurrent lending and underwriting reduces issuers' financing costs. To examine this question, we compare the yield spreads of concurrent loans and nonconcurrent loans.[12] Lower yield spreads for concurrent loans would be consistent with the existence of informational economies of scope. Second, we wish to examine if the benefits provided to concurrent issuers vary by the type of underwriter. Considering the result from the last section in which we find that investment banks discount underwriter spreads, any differences between investment bank and commercial bank pricing of concurrent loans will provide insight into how these two underwriter types compete. Therefore, we compare the yield spreads of concurrent loans in which the lender is a commercial bank with concurrent loans from investment banks.

B.1. Hand Matching

To examine pricing differences between concurrent and nonconcurrent loans, we hand-match concurrent loans to nonconcurrent loans along four dimensions: (i) loan origination date; (ii) industry; (iii) credit rating; and (iv) length of the loan. Ideally, we would like to find a nonconcurrent loan that matches the concurrent loan on all four dimensions. However, it is unlikely that we will find an exact match. Instead, for each of the 107 concurrent lending facilities in the hand-matching sample, we select the nonconcurrent loan with the closest term length, given that the nonconcurrent loan was originated between 6 months before and 6 months after the concurrent loan origination date, and the nonconcurrent borrower belongs to the same industry and has the same credit rating as the concurrent borrower.[13] Therefore, any selected nonconcurrent loan will be an exact match on two of the four dimensions (industry and credit rating) and will have a very similar term length and loan origination date.

[12] The yield spread is the rate that the borrower pays to the lender (inclusive of fees), quoted in basis points over LIBOR.

[13] We also restrict the selection of nonconcurrent loans to those that are originated between 3 months prior to and 3 months after the term facility origination date. The results are similar and are not reported. We match on the credit rating of the borrower at the loan origination date. If the bank acts rationally, it should consider the effect that the loan will have on the credit risk of the firm when determining the price and structure of the loan. Therefore, we also examine the credit rating of the firm at two quarters after the loan. In our sample of concurrent loans, only two rated borrowers had a credit rating change during the two quarters, so both measures of credit rating provide a nearly identical sample.

We examine the mean difference between concurrent and nonconcurrent loan yield spreads using three estimators.[14] The "12-month estimator" uses all matches in which the absolute value of the difference between the term lengths of the matched pair of loans is less than 12 months. The "6-month estimator" is the same as the 12-month estimator except that the difference cannot exceed 6 months. The "exact estimator" only includes matches such that each loan in a matched pair has the same term length. For all three estimators, on average, the concurrent loan yield spreads are more than 20 basis points lower than the matched nonconcurrent loan yield spreads, a significant difference at the 5% level.

B.2. Econometric Matching

There are a few problems with the hand-matching method. First, we match on only four dimensions and ignore variables that may be relevant in determining yield spread differences, such as the size of the lending facility and the type of lending facility. Second, for matching to occur, there must exist at least one nonconcurrent loan that meets these four criteria. As a result, we do not generate matches for all of the concurrent loans in our sample. To reduce these problems, we rely on econometric matching techniques that were developed by Rosenbaum and Rubin (1983) and extended by Heckman and Robb (1986) and Heckman et al. (1997, 1998).[15] In Appendix B, we provide a summary of these techniques and a detailed description of how we apply the methods to our data.

Essentially, instead of facing the difficult task of matching directly on multiple dimensions, econometric matching allows us to match nonconcurrent loans to concurrent loans based on a one-dimensional propensity score that is a function of the loans' observable characteristics. As a result, we effectively match loans based on many observable characteristics while not reducing the number of concurrent loans for which we can find matches. Furthermore, the methods take into account the fact that the characteristics of concurrent loans may differ significantly from nonconcurrent loans and ensure that such observed characteristics are not driving the results.

We choose to use econometric matching techniques instead of the alternative approach of employing a multivariate regression model because matching employs fewer restrictions than the regression approach, and many studies have confirmed that propensity score matching methods can allow for a more accurate analysis (see, e.g., Rubin (1997), Conniffe, Gash, and O'Connell (2000)). A key restriction in using multivariate regressions to study the pricing of loans is that the covariates are assumed to be linearly related to the yield spread.

[14] If multiple nonconcurrent loans share the closest term length to the nonconcurrent loan, we use the average yield spread of the nonconcurrent loans.

[15] Previous papers in economics and finance use the Heckman et al. (1997, 1998) matching methodology. McMillen and McDonald (2002) apply the method to study land valuation in a newly zoned city, while Dearden, Ferri, and Meghir (2002) and Blundell et al. (2000) use the matching methods to study the effect of education on wages. Bharath (2002) uses these methods to evaluate the agency costs of debt.

In the propensity score approach, the researcher does not need to specify the actual relation between yield spreads and the characteristics that can affect loan pricing.

In our models, the propensity score is a function of the firm's credit rating, the notional value of the loan facility, the term length of the loan, the type of lending facility, the year of the facility origination, and the firm's industry. Using propensity scores and econometric matching estimators, we calculate average yield spread differences between concurrent loans and matched nonconcurrent loans. Further, we split our sample to allow for a comparison of junk-rated concurrent loans with matched junk-rated nonconcurrent loans, and to enable concurrent loans to investment-grade-rated borrowers to be matched with nonconcurrent loans to similar investment-grade-rated borrowers. Also, we extend the methodology to capture differences between commercial bank concurrent loans and investment bank concurrent loans. We compare commercial bank concurrent loans to nonconcurrent loans by restricting the concurrent lending sample to include only commercial bank loans. Separately, we examine differences between investment bank concurrent loans and nonconcurrent loans.

B.3. Results

Each of the econometric matching estimators provides a sample of yield spread differentials, with each yield spread differential representing the discount (if negative) or premium (if positive) that a concurrent borrower pays. We calculate the sample average and standard error for the estimations and display the results in Table IV.

First, we provide evidence that is consistent with the existence of economies of scope in concurrent deals. As displayed in the first column of Table IV, all estimators indicate that concurrent loans have significantly lower yield spreads, with the average discount ranging between 9.97 and 14.81 basis points. On a $200 million, 6-year loan, a reduction of 9.97 basis points represents a present value savings of $770,000, while a 14.81-basis point reduction provides a present value savings of $1.15 million.[16]

We attempt to determine the effect of prior lending relationships on the yield spread differential between concurrent and nonconcurrent loans. For each estimator, we regress the sample of estimated yield spread differentials on a dummy variable that indicates whether the borrower of the concurrent loan has a prior lending relationship with the bank. Our results indicate that a prior lending relationship does not significantly affect the size of the discount.

Second, we find that the lower yield spreads on concurrent loans are concentrated among borrowers that have lower credit quality. The results in the second column of Table IV show that yield spreads on concurrent loans to junk-rated borrowers are discounted, on average, by between 12.10 and 15.96 basis points relative to matched nonconcurrent loans to junk-rated borrowers, and the discounts are strongly significant for all four estimators. In comparison,

[16] This calculation assumes a yearly discount rate of 15%.

Table IV

Estimated Mean Yield Spread Differences, in Basis Points

This table provides estimates of the mean difference between the yield spread (YSPREAD) of concurrent loans and other loans, using various estimators. YSPREAD is the rate that the borrower pays to the lender (inclusive of fees), quoted in basis points over LIBOR. Concurrent (other issuance period) loans are loans to the issuer of a seasoned equity offering (SEO) between 6 months prior to and 6 months after the SEO where the lender is (not, but could have been selected as) the underwriter of the SEO. We compute propensity scores, match concurrent loans with nonconcurrent loans using the propensity scores, and estimate the differences in yield spread between the two types of loans. We compute propensity scores using a probit model. The dependent variable is CONCURRENT, a dummy variable that equals 1 if the lending facility is a concurrent loan. The independent variables are as follows: RATING is the Standard & Poor's credit rating of a firm at the date of the loan identified through a numerical counterpart AAA = 1, AA = 2, A = 3, BBB = 4, BB = 5, B = 6, CCC = 7, CC = 8, C = 9; FACSIZE is the notional value of the loan facility, expressed in millions of dollars; LENGTH is the term length of the loan facility, measured in months; TYPE are dummy variables that indicate the type of loan, as classified by LPC *Dealscan*; YEAR are indicators for the loan year; and, INDUSTRY is a set of industry dummy variables based on two-digit primary SIC code. All variables are explained in detail in Appendix A. Estimators are nearest neighbor matching using n nonconcurrent loans (NEAR NEIGHBOR), and kernel-based matching techniques (GAUSSIAN and EPANECHNIKOV). The estimators are described in detail in Appendix B. For all estimations, we present the sample averages of yield spread differences. Column (1) presents results for the full sample of loans. Column (2) provides estimates for when only junk rated loans are included in the sample. Column (3) provides results for when only CB loans are included. Column (4) reports the difference between CB concurrent yield spread differences and other CB issuance period yield spread differences. We report t-ratios in parentheses, which are calculated using standard errors that are computed by bootstrapping with 50 replications. ***, **, * indicates significantly different than zero at the 1%, 5%, and 10% level, respectively.

Estimator	Concurrent and Nonconcurrent (1)	Junk-Rated Concurrent and Nonconcurrent (2)	CB Concurrent and Nonconcurrent (3)	CB Concurrent and Other CB Issuance Period (4)
NEAR NEIGHBOR	−14.811**	−13.690**	−22.713**	−28.422*
($n = 10$)	(−2.09)	(−2.23)	(−2.38)	(−2.92)
NEAR NEIGHBOR	−12.081**	−12.104**	−19.052**	−28.202**
($n = 50$)	(−2.38)	(−2.24)	(−2.31)	(−1.96)
GAUSSIAN	− 9.966*	−13.041**	−16.347**	−16.430
	(−1.93)	(−2.38)	(−2.23)	(−1.12)
EPANECHNIKOV	−14.772**	−15.959**	−21.223**	−26.409*
	(−2.27)	(−2.06)	(−2.57)	(−1.83)

we find that investment-grade borrowers do not receive significantly lower yield spreads on concurrent loans relative to matched nonconcurrent loan yield spreads. These results are consistent with economies of scope between lending and underwriting being more pronounced for issuers with lower credit ratings.

Third, we find that commercial banks provide cheaper loans to concurrent borrowers. In the third column of Table IV, we show that yield spreads on commercial bank concurrent loans are discounted by between 16.35 and 22.72 basis points relative to nonconcurrent yield spreads, and that the differences are highly significant for all four estimators. On a $200 million, 6-year loan, a concurrent borrower earns a present value savings of between $1.27 and $1.76 million through a discounted loan yield spread that is provided by its commercial bank.[17] Again, the savings provided by commercial banks are pronounced among junk-rated borrowers. While commercial banks reduce concurrent loan yield spreads, we find that yield spreads on investment bank concurrent loans are insignificantly different from those of nonconcurrent loans.[18] Concurrent lending by commercial banks, as opposed to investment banks, largely drives the difference between the yield spreads of concurrent and nonconcurrent loans.

These results, in combination with the results from Section II.A., indicate that in comparison to similar nonconcurrent issuers and borrowers, concurrent issuers pay lower underwriter spreads on the SEO and receive lower loan yield spreads. Furthermore, we find that the cost reductions are large and significant for issuers who are not investment-grade rated. These results are consistent with the existence of informational economies of scope. In addition, the concentration of savings among these firms helps explain why all deals are not concurrent, as concurrent lending is economically justified only when there are sufficient informational economies of scope.

Interestingly, we find that the form of the savings depends on the type of bank that is involved in the transaction, with investment banks providing lower underwriter spreads on the equity offering and commercial banks providing lower loan yield spreads. These savings are economically substantive. As an illustration, concurrent issuers who use investment banks receive an average savings of between $520,000 and $880,000 on a $200 million equity offering. Those who use commercial banks receive an average saving of between $1.27 and $1.76 million on a $200 million, 6-year loan.[19]

B.4. Robustness—Other Issuance Period Loans

An additional concern is that concurrent issuers are simultaneously raising equity and receiving loans and may therefore differ from other issuers. To address this concern, within the sample of nonconcurrent loans, we identify other issuance period loans, which are loans to an issuer of an SEO that are originated between 6 months prior to and 6 months after the SEO, where the lender could have been selected to underwrite the SEO but is not provided with

[17] Again, this calculation assumes a yearly discount rate of 15%.

[18] In unreported estimations, we find that investment bank concurrent loan yield spreads are insignificantly discounted between 0 and 6 basis points relative to matched nonconcurrent loan yield spreads.

[19] In addition, we attempt to determine if concurrent issuers have significantly reduced financing costs for the entire transaction. For each concurrent deal, we create an updated underwriter spread by calculating the dollar value of savings or premium for each concurrent loan, adding this value to the dollar fee for the SEO, and dividing the total amount by the principal amount of the SEO. Unreported estimations of the underwriter spread models reveal that concurrent issuers receive significant savings on the entire transaction.

underwriting responsibilities.[20] We then compare concurrent loan yield spreads with other issuance period loan yield spreads to determine whether the results in Section II.B.3. are robust.

In Section II.B.3, we show that most of the discounting of concurrent loans comes from commercial banks. Hence, we compare commercial bank concurrent loans with other commercial bank issuance period loans. Extending the previously employed methodology, we match commercial bank concurrent loans to nonconcurrent loans as well as other commercial bank issuance period loans to nonissuance period loans by computing propensity scores and calculating yield spread differences.

We compute sample averages for the concurrent loan matched pairs and the other issuance period loan matched pairs and report the mean difference in the yield spread between the two groups in the fourth column of Table IV. The results of all four estimations indicate that commercial bank concurrent loans are discounted more than other commercial bank issuance period loans. On average, concurrent loan yield spreads are less than other issuance period loan yield spreads by 16.43–28.42 basis points, and the difference is significant when using three of the four estimators. Relative to loans provided by other commercial bank underwriters around the time of the SEO issuance, the discount that is provided by commercial banks to concurrent issuers remains significant.

C. Underwriter Relationships

In Sections II.A. and II.B., we find that the issuers who participate in a concurrent deal benefit from lower financing costs in the form of lower underwriter spreads and lower loan yield spreads. Here, we examine whether underwriters benefit from concurrent lending and underwriting. Underwriters may gain if this practice helps build relationships that improve the bank's chances of capturing current or future underwriting business. Hence, we first investigate whether providing a concurrent loan significantly increases the probability that the bank wins the current equity underwriting mandate. Then we investigate if concurrent lending and underwriting increases the likelihood that the bank will receive future underwriting business from the firm, thereby increasing expected future revenues.

C.1. McFadden's Choice Model

In this section, we study the influence that concurrent lending has on the likelihood that a bank is selected as the equity underwriter. We use McFadden's (1973) choice model to capture the effect.[21]

Each issuing firm i chooses an underwriter j from a set of J underwriters. The choice of underwriter will depend on the characteristics of the issuer and attributes of the underwriter. The utility of choice j is

[20] We also extend this sample to include loans from any bank, not just those who could be selected to underwrite the SEO. The results are qualitatively similar.

[21] See Greene (2000) for a discussion of models for choices between multiple alternatives.

$$U_{ij} = \boldsymbol{\alpha}'\boldsymbol{w}_i + \boldsymbol{\beta}'\boldsymbol{x}_{ij} + \varepsilon_{ij}, \tag{1}$$

where $\boldsymbol{w}_i$ is a vector of issuer characteristics and $\boldsymbol{x}_{ij}$ is a matrix of choice attributes. If the issuing firm makes a choice j, then we assume that U_{ij} is the maximum among the J utilities. Let Y_i be a random variable that indicates the firm's choice. McFadden (1973) shows that if the J disturbances are independent and identically distributed with Weibull distribution, then

$$\Pr(Y_i = j) = \frac{\exp(\boldsymbol{\alpha}'\boldsymbol{w}_i + \boldsymbol{\beta}'\boldsymbol{x}_{ij})}{\sum_{j=1}^{J}\exp(\boldsymbol{\alpha}'\boldsymbol{w}_i + \boldsymbol{\beta}'\boldsymbol{x}_{ij})}. \tag{2}$$

We assume that each firm has 21 potential choices—each of the top-20 underwriters and a single choice of any of the underwriters that are not ranked in the top 20. Since the attributes of the potential underwriters can influence an issuer's choice, we track underwriting relationships, lending relationships, analyst coverage, and all-star analyst coverage for each of the issuer's potential choices.[22] By including this information, we more accurately control for relationship-specific and underwriter-specific factors that could affect the probability of a firm selecting an underwriter. In addition, we modify our definition of concurrent loans to include loans from potential underwriters that are originated between 6 months prior to the SEO and 6 months after the SEO. This adjustment amounts to adding the 340 "other issuance period loans" to the sample of 358 concurrent loans.[23] Technically, this modification is needed because, otherwise, concurrent lending perfectly predicts an issuer's choice of underwriter. This methodology allows us to address whether, conditional on a firm issuing seasoned equity, lending at the time of the SEO improves the probability of obtaining the underwriting business.

In our models, we assume that the relevant issuer-specific characteristics ($\boldsymbol{w_i}$) are the logarithm of the SEO principal amount, the age of the firm, the long-term debt to equity ratio of the firm in the quarter of the SEO, and the industry of the issuer. These variables are chosen to control for the differences between concurrent and nonconcurrent issuers that are shown in Table II, Panel A. For the choice-specific attributes ($\boldsymbol{x}_{ij}$), we include variables to capture concurrent lending, prior lending relationships, prior underwriting relationships, as well as the reputation of the underwriter, the level of equity analyst coverage, and the

[22] For example, even though AMC Entertainment selected Goldman Sachs to underwrite its August 1998 SEO, we capture the fact that it could have selected Morgan Stanley and that Morgan Stanley provided all-star analyst coverage for the firm. Our final data set consists of 48,321 firm–underwriter pairs (2,301 firms $\times$ 21 choices).

[23] Since multiple underwriters can be lenders on a given lending facility, the number of underwriters that provide "concurrent loans" exceeds the total number of "concurrent loans." A total of 1,154 firm–underwriter pairs have at least one loan that is originated between 6 months prior to the SEO and 6 months after the SEO, of which 201 underwriters are selected to underwrite deals. Of the 1,154 pairs, an unranked underwriter provided a loan around the equity issuance on 106 occasions.

quality of equity analyst coverage.[24] Our priors are that preexisting lending and underwriting relationships between a firm and an underwriter will increase the probability of selection. Also, we expect that the reputation of the underwriter and the level and quality of equity analyst coverage will be positively related to underwriter selection. We estimate two models—in the first model, we do not consider differences between investment banks and commercial banks; we relax this restriction in the second model.

C.2. Results

In Table V, we present the results of the underwriter selection models. In both models, the control variables have the expected signs and most are highly significant. The coefficients of all concurrent lending variables are positive and statistically significant at the 1% level. This indicates that after controlling for other factors that significantly influence underwriter selection, providing a concurrent loan increases the probability of winning the underwriting mandate, conditional on a firm issuing seasoned equity. The effect is present for both commercial and investment bank underwriters. Further, the estimation in the first column of Table V shows that having a prior lending relationship with the underwriter (both with and without a concurrent loan) increases the likelihood of selection. Overall, the results demonstrate that providing a concurrent loan, and having a prior lending relationship in general, increases a bank's expected investment banking revenues and raises the likelihood of building relationships with issuers.

D. Probability of Keeping Future Business

Lending and underwriting concurrently may also foster a durable relationship that can boost expected future revenues by increasing the likelihood that the issuer will use the bank repeatedly. Future interactions could become more likely because this practice allows the bank to generate private information that can be used in ongoing transactions with the bank, thereby providing the bank with a source for both lending and underwriting relationships.[25] In this section, we determine whether concurrent lending and underwriting enhances an underwriter's ability to cultivate relationships by examining whether those firms that participate in a concurrent deal go back to the market more frequently and switch underwriters less often than issuers who do not receive a concurrent loan.

In Table VI, we present a univariate analysis of switching probabilities. For our sample of 2,301 issuers, 37% of concurrent issuers proceed with a

[24] All underwriters that have a ranking below twentieth are modeled as a single, unranked choice. We treat this potential choice as an investment bank. If any of the variables in $\boldsymbol{x}_{ij}$ equal 1 for any of the underwriters that are ranked below twentieth, then the variable equals 1 for the unranked choice. Results are robust to excluding the unranked choice.

[25] Access to firm-specific information is well known to be a key factor in developing and maintaining lending relationships (see Ongena and Smith (2000) for a survey of the literature). Private information is also a key determinant of investment banking relationships (see, e.g., Crane and Eccles (1993)).

Table V

Multivariate Model of Underwriter Selection (McFadden's Choice Model)

This table presents the results of two models of underwriter selection. For each issue, the issuer chooses from a set of 21 lead underwriters that contains the top-20 ranked equity underwriters from the year prior to the issue and a single underwriter that represents all other underwriters. The dependent variable is 1 if the potential underwriter is selected and 0, otherwise. There are issuer-specific and choice-specific independent variables. The issuer-specific variables are as follows: LNSIZE is the logarithm of the principal amount, expressed in millions of dollars; AGE is the firm's age, measured in years; DE-LTDEBT is the long-term debt to common equity ratio in the quarter of the SEO; and, SICx are industry dummy variables, which are 1 if the firm has the corresponding one-digit SIC. The choice-specific variables are: CONCLOAN is a dummy variable that equals 1 if a potential underwriter provides a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the potential underwriter had never provided a loan to the issuer in the past; CONCPLEND is a dummy variable that equals 1 if a potential underwriter provides a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the potential underwriter provided a loan to the issuer prior to 6 months before the SEO; PRIORLEND is a dummy variable that equals 1 if a loan between the potential underwriter and the issuer was originated at any time prior to 6 months before the SEO *and* the potential underwriter does not provide a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO; PRIORUND is 1 if a potential underwriter had been the underwriter on any prior equity offering by the issuer; IB (CB) is 1 if the potential underwriter of the SEO is an investment bank (commercial bank); COVERAGE is 1 if the potential underwriter had provided an earnings per share estimate for the firm during the year prior to the SEO; ALLSTAR is 1 if COVERAGE is 1 and the analyst was ranked as an all-star by Institutional Investor magazine for the year prior to the SEO; and, RANK1 through RANK20 are 20 dummy variables, one for each ranked potential choice. All variables are explained in detail in Appendix A. In column (1), we do not consider differences between investment banks and commercial banks. In column (2), we allow for these differences by interacting both IB and CB with CONCLOAN, CONCPLEND, PRIORLEND, and PRIORUND. The choice-specific variables are interacted with the 20 rank dummy variables in order to be included in models. Estimated coefficients on the rank dummy variables and the issuer-specific variables are not reported. ***, **, * indicates significantly different than 0 at the 1%, 5%, and 10% level, respectively.

	(1)		(2)	
	Coefficient	T-ratio	Coefficient	T-ratio
CONCLOAN	1.997	10.35***		
CONCPLEND	1.574	7.23***		
PRIORLEND	0.534	3.38***		
PRIORUND	2.728	36.79***		
IB × CONCLOAN			2.086	8.32***
CB × CONCLOAN			1.838	6.83***
IB × CONCPLEND			1.920	5.79***
CB × CONCPLEND			1.439	5.54***
IB × PRIORLEND			0.883	4.37***
CB × PRIORLEND			0.141	0.58
IB × PRIORUND			2.898	33.02***
CB × PRIORUND			2.244	15.31***
IB			−0.107	−1.14
COVERAGE	1.618	19.57***	1.655	19.89***
ALLSTAR	0.582	4.83***	0.559	4.62***
Psuedo R^2	0.4341		0.4362	
Log likelihood	5,053.52		5,078.83	

Table VI

Univariate Analysis of Keeping the Same Underwriter in a Subsequent SEO

This table summarizes the probability that an issuer will proceed with a subsequent seasoned equity offering (SEO) and, if so, the probability that the issuer will keep the underwriter, based on whether the initial SEO was a concurrent deal. A concurrent deal is any SEO in which the underwriter provided a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO. The underwriter is an IB (CB) if the parent or holding company of the underwriter is an investment bank (commercial bank) at the time of the SEO. Panel A provides a full sample analysis. Panel B examines those SEOs in which the underwriter is an investment bank. Panel C examines those SEOs in which the underwriter is a commercial bank. p-Values for the difference in proportions are provided in the last column. ***, **, and * indicate significantly different than 0 at the 1%, 5%, and 10% level, respectively.

	Concurrent Deals	Nonconcurrent Deals	Proportion Test p-Value
Panel A: Full Sample			
# in sample	201	2,100	
# that repeat	74	462	
% of sample that repeat	36.82%	22.00%	0.0000***
# keep same underwriter	42	207	
% of repeaters that keep same underwriter	56.76%	44.81%	0.0556*
Panel B: Underwriter Is an IB			
# in sample	110	1,509	
# that repeat	43	347	
% of sample that repeat	39.09%	23.00%	0.0001***
# keep same underwriter	28	148	
% of repeaters that keep same underwriter	65.12%	42.65%	0.0049***
Panel C: Underwriter Is a CB			
# in sample	91	591	
# that repeat	31	115	
% of sample that repeat	34.07%	19.46%	0.0018***
# keep same underwriter	14	59	
% of repeaters that keep same underwriter	45.16%	51.30%	0.5162

subsequent equity offering, while only 22% of nonconcurrent issuers go back to the equity market.[26] Of those firms that have a follow-up equity offering, 57% of concurrent issuers and 45% of nonconcurrent issuers keep the same underwriter, a significant difference at the 10% level. However, there is a disparity between investment bank and commercial bank underwriters. While a prior concurrent deal significantly increases the probability of retaining future business for investment banks, the effect is not present for commercial banks. This result indicates that commercial banks may not be able to leverage their

[26] We examine subsequent SEOs that took place before March 31, 2002. Extending the sample end date allows issuers from the latter part of the sample to potentially reissue.

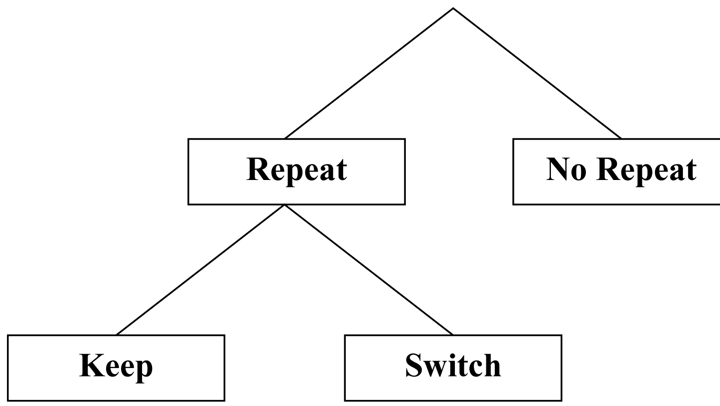


Figure 1. Nesting structure. This figure presents the nesting structure for the nested logit model of keeping the same underwriter in a subsequent SEO. Each issuer has a first-level choice of reissuing ("REPEAT") or not reissuing ("NO REPEAT"). If the issuer decides to reissue, the issuer has a second-level choice of keeping the underwriter of the current SEO ("KEEP") or switching to a new underwriter ("SWITCH") in the subsequent offering.

practice of concurrently lending and underwriting into extended underwriting relationships.

D.1. Nested Logit Model

To determine whether these results withstand a multivariate specification, we use a nested logit model. As shown in Figure 1, we assume that each issuer makes a two-stage decision. First, the issuer decides if it will proceed with a subsequent SEO or if it will not issue again. Second, if the issuer chooses to issue again, then it can keep the same underwriter or switch to a new underwriter.

Following Maddala (1983), let k index the first-level alternative and l the second-level alternative.[27] Also, let $\boldsymbol{Y}_{kl}$ and $\boldsymbol{Z}_k$ be vectors of explanatory variables specific to the categories (k, l) and (k), respectively. Then each issuer will have a utility U_{kl} for alternative (k, l) that is a function of the explanatory variables. We set $U_{kl} = \alpha'\mathbf{Y}_{kl} + \beta'\mathbf{Z}_k + \varepsilon_{kl}$, and then the probability of choosing l, conditional on first choosing k, is

$$\mathrm{Pr}_{l\,|k} = \frac{\exp(\boldsymbol{\alpha}'\boldsymbol{Y}_{kl})}{\sum_{l=1}^{L}\exp(\boldsymbol{\alpha}'Y_{kl})}. \tag{3}$$

Define the inclusive values for category (k) as

$$\mathrm{IV}_{\boldsymbol{k}} = \ln\left(\sum_{l=1}^{L}\exp(\boldsymbol{\alpha}'\mathbf{Y}_{kl})\right), \tag{4}$$

which leaves us with the probability of choosing k as

[27] For our model, k can be "Repeat" or "No Repeat" while l can be "Keep" or "Switch."

$$\Pr_k = \frac{\exp(\beta'\mathbf{Z}_k + \tau_k \mathrm{IV}_k)}{\sum_{k=1}^{K} \exp(\beta'\mathbf{Z}_k + \tau_k \mathrm{IV}_k)}. \tag{5}$$

In our models, we assume that the variables that only affect the decision to reissue ($\boldsymbol{Z}_k$) are the logarithm of the SEO principal amount, the age of the firm, the long-term debt to equity ratio of the firm in the quarter of the SEO, and the industry of the issuer. For the variables that affect both the decision to reissue and the decision to keep or switch underwriters ($\boldsymbol{Y}_{kl}$), we include variables to capture concurrent lending, prior lending relationships, prior underwriting relationships, as well as differences between the original underwriter and the subsequent underwriter in the level and quality of equity analyst coverage and underwriter ranking. We expect that prior lending and underwriting relationships will be positively related to retaining future business. Also, previous papers indicate that firms will be more likely to switch to an underwriter who has higher quality equity analyst coverage and is ranked above the original underwriter (see, e.g., Krigman et al. (2001), Fernando et al. (2005)). As in the previous section, we estimate one model in which we do not consider differences between investment banks and commercial banks and a second model in which we relax this restriction. Based on the univariate results, we expect a previous concurrent deal with an investment bank underwriter to increase the probability that the investment bank retains future underwriting business. We also expect that a previous concurrent deal with a commercial bank will not significantly affect the probability that the bank can retain equity underwriting business in the future.

D.2. Results

In Table VII, we present the results of the nested logit models. The base case category is that the issuer does not have a subsequent equity offering, so variables that are interacted with KEEP provide the effects of choosing to reissue and keep the same underwriter instead of not reissuing at all. We also determine the effect of the variables on keeping the same underwriter instead of switching to a new underwriter through *t*-tests for differences between keeping and switching.

In the first column of Table VII, we present the results of the model in which we do not consider differences between investment banks and commercial banks. We find that a prior concurrent deal increases the probability of an issuer choosing to reissue and keep the same lead underwriter relative to not reissuing. The *t*-tests for differences between keeping and switching indicate that a previous concurrent deal also increases the probability of keeping an underwriter instead of switching to a new underwriter, although this result is insignificant. Furthermore, we find that prior lending relationships (both with and without a concurrent loan) increase the probability of an issuer choosing to keep the same lead underwriter. These results highlight the importance of lending in generating future investment banking business.

The second column of Table VII shows the results for the case in which we allow the coefficients to reflect disparities between investment banks and commercial banks. We find that a prior concurrent deal (without the existence of a prior lending relationship) with an investment bank significantly increases the

Table VII

Multivariate Model of Keeping the Same Underwriter in a Subsequent SEO

In this table, we present results of two nested logit models of the probability of keeping or switching underwriters in a subsequent seasoned equity offering (SEO). Let the alternatives of "Repeat" and "Not Repeat" belong to the upper nest and the alternatives of "Keep" and "Switch" belong to the lower nest. The dependent variable indicates whether the issuer decides to either (i) "Not Repeat," (ii) "Repeat" and "Keep", or (iii) "Repeat" and "Switch." The following independent variables affect the decision in the upper nest only: LNSIZE is the logarithm of the original SEO principal amount, expressed in millions of dollars; AGE is the firm's age, measured in years; DE-LTDEBT is the long-term debt to common equity ratio in the quarter of the original SEO; and, SICx are industry dummy variables, which are one if the firm has the corresponding one-digit SIC. The following independent variables affect both the upper nest and lower nest decisions: CONCLOAN is a dummy variable that equals 1 if the underwriter provides a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the underwriter had never provided a loan to the issuer in the past; CONCPLEND is a dummy variable that equals 1 if the underwriter provides a loan to the issuer between 6 months prior to the original SEO and 6 months after the original SEO *and* the underwriter provided a loan to the issuer prior to 6 months before the SEO; PRIORLEND is a dummy variable that equals 1 if a loan between the underwriter and the issuer was originated at any time prior to 6 months before the SEO *and* the underwriter does not provide a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO; PRIORUND is 1 if the underwriter had been the underwriter on any equity offering by the issuer prior to the original SEO; IB (CB) is 1 if the underwriter of the original SEO is an investment bank (commercial bank); CNGCOV is the difference between the coverage provided by the subsequent underwriter and the original underwriter; CNGSTAR is the difference between the all-star coverage provided by the subsequent underwriter and the original underwriter; CNGRANK is the difference between the subsequent underwriter's ranking and the original underwriter's ranking; and, KEEP and SWITCH are choice-specific dummy variables. All variables are explained in detail in Appendix A. In column (1), we do not consider differences between investment banks and commercial banks. In column (2), we allow for these differences by interacting both IB and CB with CONCLOAN, CONCPLEND, PRIORLEND, and PRIORUND. CONCLOAN, PRIORUND, PRIORLEND, and IB are interacted with KEEP and SWITCH in order to be included in the models. LNSIZE, AGE, DE-LTDEBT, and SICx are interacted with REPEAT in order to be included in the models. Estimated coefficients for the industry variables (SICx) are not reported. ***, **, * indicates significantly different than zero at the 1%, 5%, and 10% level, respectively.

	(1)		(2)	
	Coefficient	T-ratio	Coefficient	T-ratio
Variables That Affect the Choice of "REPEAT" or "NO REPEAT"				
REPEAT × LNSIZE	0.124	2.29**	0.139	2.55**
REPEAT × AGE	0.003	1.20	0.002	0.74
REPEAT × DE-LTDEBT	0.010	1.05	0.010	1.08
Variables That Affect the Choice of "NO REPEAT," "(REPEAT, KEEP)," or "(REPEAT, SWITCH)"				
Concurrent lending/no prior lending relationship				
KEEP × CONCLOAN	0.434	2.52**		
KEEP × IB × CONCLOAN			0.727	3.70***
KEEP × CB × CONCLOAN			−0.188	−0.44
SWITCH × CONCLOAN	0.095	0.45		
SWITCH × IB × CONCLOAN			−0.083	−0.27
SWITCH × CB × CONCLOAN			0.478	1.74*
Concurrent lending with prior lending relationship				
KEEP × CONCPLEND	0.380	1.87*		
KEEP × IB × CONCPLEND			0.071	0.19
KEEP × CB × CONCPLEND			0.603	2.23**
SWITCH × CONCPLEND	−0.008	−0.03		
SWITCH × IB × CONCPLEND			0.014	0.04
SWITCH × CB × CONCPLEND			0.125	0.36

(*continued*)

Table VII—*Continued*

	(1)		(2)	
	Coefficient	*T*-ratio	Coefficient	*T*-ratio
Prior lending relationship / no concurrent lending				
KEEP × PRIORLEND	0.320	1.71*		
KEEP × IB × PRIORLEND			0.161	0.64
KEEP × CB × PRIORLEND			0.632	2.18**
SWITCH × PRIORLEND	0.018	0.08		
SWITCH × IB × PRIORLEND			0.053	0.19
SWITCH × CB × PRIORLEND			0.025	0.05
Prior underwriting relationship				
KEEP × PRIORUND	0.282	2.77***		
KEEP × IB × PRIORUND			0.159	1.31
KEEP × CB × PRIORUND			0.557	2.91***
SWITCH × PRIORUND	−0.112	−1.08		
SWITCH × IB × PRIORUND			−0.188	−1.53
SWITCH × CB × PRIORUND			0.072	0.35
Coverage and reputation				
SWITCH × CNGCOV	0.120	0.62	0.097	0.49
SWITCH × CNGSTAR	0.737	2.36**	0.704	2.26**
SWITCH × CNGRANK	0.146	7.72***	0.146	7.55***
Bank classification and constants				
KEEP × IB			0.250	1.38
SWITCH × IB			0.312	1.85*
KEEP	−1.494	−8.41***	−1.730	−7.14***
SWITCH	−1.303	−8.32***	−1.582	−6.78***
IV(REPEAT)	2.490	6.83***	2.441	6.68***
LR test of homoskedasticity (IV(Repeat) = 1)	34.97***		32.30***	
Log likelihood	1,315.01		1,301.27	
t-Tests for Differences between Keeping and Switching				
KEEP × CONCLOAN–SWITCH × CONCLOAN	0.339	1.05		
KEEP × IB × CONCLOAN–SWITCH × IB × CONCLOAN			0.810	1.92*
KEEP × CB × CONCLOAN–SWITCH × CB × CONCLOAN			−0.667	−1.10
KEEP × CONCPLEND–SWITCH × CONCPLEND	0.388	1.00		
KEEP × IB × CONCPLEND–SWITCH × IB × CONCPLEND			0.057	0.09
KEEP × CB × CONCPLEND–SWITCH × CB × CONCPLEND			0.478	0.93
KEEP × PRIORLEND–SWITCH × PRIORLEND	0.303	0.82		
KEEP × IB × PRIORLEND–SWITCH × IB × PRIORLEND			0.108	0.23
KEEP × CB × PRIORLEND–SWITCH × CB × PRIORLEND			0.608	0.97
KEEP × PRIORUND–SWITCH × PRIORUND	0.394	2.21**		
KEEP × IB × PRIORUND–SWITCH × IB × PRIORUND			0.347	1.62
KEEP × CB × PRIORUND–SWITCH × CB × PRIORUND			0.485	1.44

probability of keeping the same underwriter in the subsequent equity offering. The results indicate that for commercial bank underwriters, a concurrent deal does not significantly affect the probability that an underwriter will keep the same underwriter instead of switching to a new underwriter in the subsequent equity offering. These results are consistent with the univariate statistics in Table VI.

Combined with our previous findings, we find that investment banks discount underwriter spreads and that concurrent deals increase the probability of retaining future underwriting business from the firm. Commercial banks, on the other hand, discount loan yield spreads that can help establish lending relationships that are well known to lead to other fee-based lending business (for some recent evidence, see Bharath et al. (2004)). Therefore, the results are consistent with each type of underwriter competing more aggressively in its area of expertise and in the area in which it is more likely to generate future business.

III. Conclusion

We use a unique data set drawn from multiple data sources and augmented by hand-collected data to examine when financial institutions concurrently lend and underwrite, and to analyze the effect on firms' financing costs and firm–underwriter relationships. We find evidence that is consistent with concurrent lending and underwriting occurring when there are large potential efficiency gains that arise due to informational economies of scope from combining lending and equity underwriting. This is supported by the preponderance of concurrent deals involving highly leveraged and noninvestment-grade issuers and the substantial benefits that concurrent deals bring to such issuers. For issuers, these benefits come in the form of lower financing costs, as concurrent issuers receive a lower underwriter fee for the equity offering and a discounted yield spread on the concurrent loan. The cost reductions are large and significant for issuers who are noninvestment-grade rated, where the expected informational economies of scope are sizeable. Interestingly, the benefit that an issuer receives varies by the type of underwriter involved in the transaction. Investment banks offer reduced underwriter spreads on concurrent SEOs, while commercial banks offer discounted loan yield spreads. This is consistent with each type of underwriter competing more aggressively in its area of expertise. In addition to benefiting issuers, concurrent lending and underwriting produces gains for underwriters. We find that providing a concurrent loan increases the likelihood of receiving the current equity underwriting business, and also helps generate other business from the issuers, with investment bank underwriters more likely to receive future equity underwriting mandates from concurrent issuers. These results are consistent with concurrent loans being used to help build ongoing, durable relationships that increase an underwriter's expected revenues. This study also highlights the importance of prior lending relationships, in general. Issuers with prior lending relationships receive significantly lower underwriter spreads and, for the underwriter, a prior lending relationship increases the likelihood of receiving underwriting business. Our findings of substantial benefits to issuers and underwriters from combining lending with underwriting indicates that concurrent lending is likely to continue into the future and that lending will influence the pricing of financial products and services as well as remain an important factor in determining firm–underwriter pairings.

Appendix A

Detailed Descriptions of the Variables

Underwriter Spread Regressions (Section II.A.)

USPREAD: The underwriter spread, which is the compensation paid to the underwriter for selling the firm's security issue, expressed as a percent of the capital raised.

CONCLOAN: A dummy variable that equals 1 if the underwriter provided a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the underwriter had never provided a loan to the issuer in the past.

CONCPLEND: A dummy variable that equals 1 if the underwriter provided a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the underwriter provided a loan to the issuer prior to 6 months before the SEO.

PRIORLEND: A dummy variable that equals 1 if a loan between the underwriter and the issuer was originated at any time prior to 6 months before the SEO *and* the underwriter does not provide a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO.

PRIORUND: A dummy variable that equals 1 if the underwriter had been the underwriter on any prior equity offering by the issuer.

IB: A dummy variable that equals 1 if the parent/holding company of the underwriter at the time of the issue is an investment bank.

CB: A dummy variable that equals 1 if the parent/holding company of the underwriter at the time of the issue is a commercial bank.

(1/SEOSIZE): The inverse of the principal amount of the offering (in millions of dollars). This variable captures the fixed cost component of underwriter spreads.

(SEOSIZE / MKTCAP): The principal amount of the offering divided by the market capitalization of the issuer at the date of the SEO. This variable captures the variable cost component of underwriting spreads.

VOL: The daily standard deviation of the issuer's common stock rate of return over the 220 trading days ending 40 days before the offering.

MKTACT: The dollar volume of issuance by firms whose SIC codes are not six in the U.S. seasoned equity market during the 3 months prior to the SEO date.

Propensity Score / The Pricing of Concurrent Loans (Section II.B.)

YSPREAD: The yield spread of the loan, measured as the rate the borrower pays to the lender, quoted in basis points over the London Interbank offer rate (LIBOR). We use the *DealScan* item "all-in spread drawn," which adds the spread of the loan to any fees that have to be paid back to the bank.

CONCURRENT: A dummy variable that equals 1 if the lending facility is to an issuer of an SEO and is originated between 6 months prior to and 6 months after the SEO and the lender is selected to underwrite the SEO.

ISSPERLOAN: A dummy variable that equals 1 if the lending facility is to an issuer of an SEO and is originated between 6 months prior to and 6 months after the SEO, where the lender could have been selected to underwrite the SEO but is not provided with underwriting responsibilities.

RATING: A variable that provides the Standard & Poor's credit rating of the firm at the date of the lending facility. Each rating is given a numerical counterpart: AAA = 1, AA = 2, A = 3, BBB = 4, BB = 5, B = 6, CCC = 7, CC = 8, and C = 9.

FACSIZE: The notional value of the loan facility between the lender and the borrower, expressed in millions of dollars.

LENGTH: The term length of the loan, measured as the difference between the term facility active date and the term facility expiration date, measured in months.

TYPE: Dummy variables that correspond to the type of lending facility. The dummy variables indicate whether the facility is a term loan, 364-day facility, revolving line of credit, or other type.

YEAR: Dummy variables that correspond to the year of the origination date of the lending facility.

INDUSTRY: Dummy variables that equal 1 if the borrower is in the corresponding two-digit SIC group.

(*continued*)

Appendix A—*Continued*

McFadden Choice Model/Underwriter Relationships (Section II.C.)

CONCLOAN: A dummy variable that equals 1 if a potential underwriter provided a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the potential underwriter had never provided a loan to the issuer in the past.

CONCPLEND: A dummy variable that equals 1 if the potential underwriter provided a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO *and* the potential underwriter provided a loan to the issuer prior to 6 months before the SEO.

PRIORLEND: A dummy variable that equals 1 if a loan between the potential underwriter and the issuer was originated at any time prior to 6 months before the SEO *and* the potential underwriter does not provide a loan to the issuer between 6 months prior to the SEO and 6 months after the SEO.

PRIORUND: A dummy variable that equals 1 if the potential underwriter had been the underwriter on any prior equity offering by the issuer.

COVERAGE: A dummy variable that equals 1 if the potential underwriter provided an earnings per share estimate for the firm during the year prior to the SEO.

ALLSTAR: A dummy variable that equals 1 if COVERAGE is one and the analyst was ranked as an all-star by *Institutional Investor* magazine for the year prior to the SEO.

RANK: We compute each underwriter's yearly SEO market share by adding the principal amounts of all SEOs in which the bank was an underwriter and dividing this total by the principal amounts of all SEOs during the year. To avoid potential simultaneity problems, we rank the underwriters on a yearly basis, based on the market share in the previous year. If a merger between underwriters occurred during the year, we use the combined market share of the underwriters. The top-ranked underwriter is given a score of 20, the second ranked underwriter is 19, and so on. Underwriters not ranked in the top 20 are given a score of 0.

Nested Logit Model/Keeping Future Business (Section II.D.)

CONCLOAN: A dummy variable that equals 1 if the underwriter provided a loan to the issuer between 6 months prior to the original SEO and 6 months after the original SEO *and* the underwriter had never provided a loan to the issuer in the past.

CONCPLEND: A dummy variable that equals 1 if the underwriter provided a loan to the issuer between 6 months prior to the original SEO and 6 months after the original SEO *and* the underwriter provided a loan to the issuer prior to 6 months before the original SEO.

PRIORLEND: A dummy variable that equals 1 if a loan between the underwriter and the issuer was originated at any time prior to 6 months before the original SEO *and* the underwriter does not provide a loan to the issuer between 6 months prior to the original SEO and 6 months after the original SEO.

PRIORUND: A dummy variable that equals 1, if the underwriter had been the underwriter on any equity offering prior to the original SEO by the issuer.

REPEAT: A dummy variable that equals 1 if the issuer has a subsequent offering.

KEEP: A dummy variable that equals 1 if the issuer keeps the same underwriter in the subsequent offering.

SWITCH: A dummy variable that equals 1 if the issuer switches underwriters in the subsequent offering.

CNGCOV: For "switchers," the difference between the coverage provided by the new underwriter and the original underwriter during the year prior to the subsequent SEO. The variable can take on the values of -1, 0, or 1. By definition, for all nonrepeaters and keepers, it has a value of 0.

CNGSTAR: For "switchers," the difference between the all-star coverage provided by the new underwriter and the original underwriter during the year prior to the subsequent SEO. The variable can take on the values of -1, 0, or 1. By definition, for all nonrepeaters and keepers, it has a value of 0.

(*continued*)

Appendix A—*Continued*

CNGRANK: For "switchers," the difference between the subsequent underwriter's ranking in the year before the subsequent issue date and the original underwriter's ranking in the year before the subsequent issue date. For keepers and nonrepeaters, the variable is 0.
IB: A dummy variable that equals 1 if the parent/holding company of the potential underwriter at the time of the issue is an investment bank.
CB: A dummy variable that equals 1 if the parent/holding company of the potential underwriter at the time of the issue is a commercial bank.

Control Variables

LNSIZE: The logarithm of the principal amount of the offering.
DE-LTDEBT: The long-term debt to equity ratio in the quarter of the SEO.
AGE: The firm's age, measured as the difference between the SEO date and the incorporation date, expressed in years.
SICx: Dummy variables that equal 1 if the issuer is in the corresponding one-digit SIC group.
IGRADE: A dummy variable that equals 1 if the issuer is rated AAA, AA, A, or BBB in the quarter of the SEO by Standard & Poor's.
JUNK: A dummy variable that equals 1, if the issuer is rated BB, B, CCC, CC, or C in the quarter of the SEO by Standard & Poor's.

Appendix B: Econometric Matching Methodology

Econometric matching techniques were developed by Rosenbaum and Rubin (1983) and extended by Heckman and Robb (1986) and Heckman et al. (1997, 1998). Below, we provide a summary of their results and how we apply these methods to our data.

Let $D = 1$ if the loan is a concurrent loan, and let $D = 0$ if the loan is a nonconcurrent loan. In principle, the i^{th} concurrent loan has its observed "concurrent" yield spread Y_{1i} and another yield spread Y_{0i} that would result if it were a nonconcurrent loan. To determine the average effect of concurrent lending on yield spreads, one would calculate the mean difference between Y_{1i} and Y_{0i} for all concurrent loans. However, since we do not observe Y_{0i} for our sample of concurrent loans, we have a missing data problem that cannot be solved at the level of the individual, so we reformulate the problem at the population level. We focus on $E(Y_1 - Y_0 \mid D = 1, X)$, the mean effect of the difference between concurrent loans and nonconcurrent loans with characteristics X. While the mean $E(Y_1 \mid D = 1, X)$ can be identified from data on concurrent loans, some assumptions must be made to identify the unobservable counterfactual mean, $E(Y_0 \mid D = 1, X)$. The observable outcome of nonconcurrent loans $E(Y_0 \mid D = 0, X)$ can be used to approximate $E(Y_0 \mid D = 1, X)$. The selection bias that arises from this approximation is $B(X) = E(Y_0 \mid D = 1, X) - E(Y_0 \mid D = 0, X)$.

We use a method of matching that solves the evaluation problem. Following Heckman and Robb (1986), we assume that all relevant differences between concurrent loans and nonconcurrent loans are captured by their observable characteristics X. Let $(Y_0, Y_1) \perp D \mid X$ denote the statistical independence of $(Y_0,$

Y_1) and D conditional on X. Rosenbaum and Rubin (1983) establish that when $(Y_0, Y_1) \perp D \,|\, X$ and $0 < P(D = 1 \,|\, X) < 1$ (which are referred to as the strong ignorability conditions), then $(Y_0, Y_1) \perp D \,|\, P(D = 1 \,|\, X)$. While it is often difficult to match on high dimension X, this result allows us to match based on the one-dimensional $P(D = 1 \,|\, X)$ alone. The propensity score, $P(D = 1 \,|\, X)$, can be estimated using probit or logit models. Heckman et al. (1998) extend this result by showing that the strong ignorability conditions are overly restrictive for the estimation of $E(Y_1 - Y_0 \,|\, D = 1, X)$. Instead, a weaker mean independence condition $E(Y_0 \,|\, D = 1, P(D = 1 \,|\, X)) = E(Y_0 \,|\, D = 0, P(D = 1 \,|\, X))$ is all that is required.

To implement econometric matching, we compute propensity scores for each of the concurrent loans and nonconcurrent loans. There may be loans that have propensity scores that are outside of the common support of concurrent loan and nonconcurrent loan propensity scores. Using loans that fall outside of the common support can substantially bias the results (see, e.g., Heckman et al. (1997)). As a result, we remove all loans that are outside of the common propensity score support.

We use two classes of propensity score matching estimators: (i) nearest neighbor matching and (ii) kernel-based matching.[28] Let Y_{1i} be the yield spread of a concurrent loan, Y_{0j} be the yield spread of a nonconcurrent loan, and let $\bar{Y}_{0i}^{z}$ represent the weighted average of yield spreads of the nonconcurrent loans using estimator z that is matched with Y_{1i}. We compute the sample average of yield spread differences, $Y_{1i} - \bar{Y}_{0i}^{z}$.

For each concurrent loan, the nearest neighbor matching estimator chooses the n nonconcurrent loans with closest propensity scores to the concurrent loan propensity score. The estimator computes the arithmetic average of the yield spreads of these n nonconcurrent loans. For each Y_{1i}, we match

$$\bar{Y}_{0i}^{NN} = \frac{1}{n} \sum_{j \in N(i)} Y_{0j}, \tag{A1}$$

where $N(i)$ is the set of nonconcurrent loans that are nearest neighbors. We set $n = 10$ and $n = 50$.

The kernel estimators construct matches for each concurrent loan by using weighted averages of yield spreads of multiple nonconcurrent loans. If weights from a typical symmetric, nonnegative, unimodal kernel $K(\bullet)$ are used, then the kernel places higher weight on loans close in terms of $P(D = 1 \,|\, X)$ and lower or zero weight on more distant observations. Let

$$K_{ij} = K\left(\frac{P(X_{1i}) - P(X_{0j})}{h}\right), \tag{A2}$$

where h is a fixed bandwidth and $P(X) = P(D = 1 \,|\, X)$. For each Y_{1i}, we match a corresponding $\bar{Y}_{0i}^{K}$, where

[28] The propensity score matching methods are discussed in greater detail in Heckman et al. (1997, 1998).

$$\bar{Y}_{0i}^{K} = \frac{\sum_j K_{ij} Y_{0j}}{\sum_j K_{ij}}. \tag{A3}$$

We use two different kernels to compute $\bar{Y}_{0i}^{K}$. The Gaussian kernel uses all nonconcurrent loans, while the Epanechnikov kernel only uses nonconcurrent loans with a propensity score $P(X_{0j})$ that falls within the fixed bandwidth h of $P(X_{1i})$. We set $h = 0.01$. As a robustness check, we also try different values of h and obtain similar results.

To determine whether econometric matching is a viable method of evaluation, Heckman et al. (1997, 1998) identify four features of the data and establish matching techniques that can substantially reduce bias: (i) participants and controls have the same distributions of unobserved attributes; (ii) participants have the same distributions of observed attributes; (iii) outcomes and characteristics are measured in the same way for both groups; and, (iv) participants and controls are from the same economic environment. Items (iii) and (iv) are met very well in this study because the loan yield spreads and other loan characteristics are measured in the same way for both concurrent and nonconcurrent loans, and the nonconcurrent loans are from the same time period as the concurrent loans. To satisfy condition (ii), we use loan characteristics to match concurrent loans to nonconcurrent loans. Feature (i) cannot be achieved in a nonexperimental evaluation. However, Heckman et al. (1997) note that feature (i) is only a small part of bias in their experimental study. Thus, the method of matching nonconcurrent loans to concurrent loans can produce a viable estimate of the difference between nonconcurrent loan and concurrent loan yield spreads.

REFERENCES

Altinkilic, Oya, and Robert Hansen, 2000, Are there economies of scale in underwriting fees? Evidence of rising external financing costs, *Review of Financial Studies* 13, 191–218.

Ang, James, and Terry Richardson, 1994, The underwriting experience of commercial bank affiliates prior to the Glass-Steagall Act: A re-examination of evidence for passage of the act, *Journal of Banking and Finance* 18, 351–395.

Baker, Wayne, 1990, Market networks and corporate behavior, *American Journal of Sociology* 96, 589–625.

Benston, George, 1990, *The Separation of Commercial and Investment Banking* (Oxford University Press, New York).

Benzoni, Luca, and Carola Schenone, 2004, Conflict of interest or certification? Evidence from IPOs underwritten by the firm's relationship bank, Working paper, University of Minnesota.

Berger, Allen, Rebecca Demsetz, and Philip Strahan, 1999, The consolidation of the financial services industry: Causes, consequences, and implications for the future, *Journal of Banking and Finance* 23, 135–194.

Best, Ronald, and Hang Zhang, 1993, Alternative information sources and the information content of bank loans, *Journal of Finance* 48, 1507–1522.

Bharath, Sreedhar, 2002, Agency costs, bank specialness and renegotiation, Working paper, New York University.

Bharath, Sreedhar, Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan, 2004, So what do I get? The bank's view of lending relationships, Working paper, New York University.

Billett, Matthew, Mark Flannery, and Jon Garfinkel, 1995, The effect of lender identity on a borrowing firm's equity return, *Journal of Finance* 50, 699–718.

Blundell, Robert, Lorraine Dearden, Alissa Goodman, and Howard Reed, 2000, The returns to higher education in Britain: Evidence from a British cohort, *Economic Journal* 110, F82–F99.

Boot, Arnoud, and Anjan Thakor, 1997, Banking scope and financial innovation, *Review of Financial Studies* 10, 1099–1131.

Booth, James, and Richard Smith, 1986, Capital raising, underwriting and the certification hypothesis, *Journal of Financial Economics* 15, 261–281.

Carter, Richard, and Steven Manaster, 1990, Initial public offerings and underwriter reputation, *Journal of Finance* 45, 1045–1067.

Chaplinsky, Susan, and Gayle Erwin, 2001, Commercial banks as equity underwriters, Working paper, University of Virginia.

Clarke, Jonathan, Ajay Khorana, Ajay Patel, and P. Raghavendra Rau, 2003, Do all-star analysts drive investment banking deal flow? Working paper, Wake Forest University.

Conniffe, Denis, Vanessa Gash, and Philip J. O'Connell, 2000, Evaluating state programmes: "Natural experiments" and propensity scores, *The Economic and Social Review* 31, 283–308.

Corwin, Shane, and Paul Schultz, 2005, The role of IPO underwriting syndicates: Pricing, information production, and underwriter competition, *Journal of Finance* 60, 443–486.

Crane, Robert, and Dwight Eccles, 1993, Customer relationships in the 1990s, in Samuel Hayes, ed.: *Financial Services: Perspectives and Challenges* (Harvard Business School Press, Boston).

Dearden, Lorraine, Javier Ferri, and Costa Meghir, 2002, The effect of school quality on educational attainment and wages, *Review of Economics and Statistics* 84, 1–20.

Dingell, John, 2002, Letter to FRB and OCC re: "Pay to play" practices, Jul 11 http://www.house.gov/commerce_democrats/press/107/tr179.shtml.

Fernando, Chitru S., Vladimir A. Gatchev, and Paul A. Spindt, 2005, Wanna dance? How firms and underwriters choose each other, *Journal of Finance*.

Gande, Amar, Manju Puri, Anthony Saunders, and Ingo Walter, 1997, Bank underwriting of debt securities: Modern evidence, *Review of Financial Studies* 10, 1175–1202.

Greene, William, 2000, *Econometric Analysis*, 4th ed. (Prentice-Hall, New Jersey).

Heckman, James, Hidehiko Ichimura, and Petra Todd, 1997, Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *Review of Economic Studies* 64, 605–654.

Heckman, James, Hidehiko Ichimura, and Petra Todd, 1998, Matching as an econometric evaluation estimator, *Review of Economic Studies* 65, 261–294.

Heckman, James, and Richard Robb, 1986, Alternative method for solving the problem of selection bias in evaluating the impact of treatments on outcomes, in Howard Wainer, ed.: *Drawing Inferences from Self-Selected Samples* (Springer-Verlag, New York).

Hubbard, R. Glenn, Kenneth Kuttner, and Darius Palia, 2002, Are there bank effects in borrowers' costs of funds? Evidence from a matched sample of borrowers and banks, *Journal of Business* 75, 559–582.

James, Christopher, 1987, Some evidence on the uniqueness of bank loans, *Journal of Financial Economics* 19, 217–235.

James, Christopher, 1992, Relationship-specific assets and the pricing of underwriter services, *Journal of Finance* 47, 1865–1885.

James, Christopher, and David C. Smith, 2000, Are banks still special? New evidence on their role in the corporate capital-raising process, *Journal of Applied Corporate Finance* 13, 395–422.

Kanatas, George, and Jianping Qi, 1998, Underwriting by commercial banks: Incentive conflicts, scope economies, and project quality, *Journal of Money, Credit, and Banking* 30, 119–133.

Kanatas, George, and Jianping Qi, 2003, Integration of lending and underwriting: Implications of scope economies, *Journal of Finance* 58, 1167–1191.

Krigman, Laurie, Wayne Shaw, and Kent Womack, 2001, Why do firms switch underwriters?, *Journal of Financial Economics* 60, 245–284.

Kroszner, Randall, and Raghuram Rajan, 1994, Is the Glass-Steagall Act justified? A study of the U.S. experience with universal banking before 1933, *American Economic Review* 84, 810–832.

Ljungqvist, Alexander, Felicia Marston, and William J. Wilhelm, 2004, Competing for securities underwriting mandates: Banking relationships and analyst recommendations, *Journal of Finance* (forthcoming).

Lummer, Scott, and John McConnell, 1989, Further evidence on bank lending process and capital market response to bank loan agreements, *Journal of Financial Economics* 25, 52–63.

Maddala, G. S., 1983, *Limited Dependent and Qualitative Variables in Econometrics* (Cambridge University Press, New York).

McFadden, Daniel, 1973, Conditional logit analysis of qualitative choice behavior, in Paul Zarembka, ed.: *Frontiers in Econometrics* (Academic Press, New York).

McMillen, Daniel P., and John F. McDonald, 2002, Land values in a newly zoned city, *Review of Economics and Statistics* 84, 62–72.

Michaels, Adrian, and Gary Silverman, 2003, Fed seeks to clarify the rules on "tying", *Financial Times*, August 26, 25.

Ongena, Steven, and David Smith, 2000, Bank relationships: A review, in Patrick Harker and Stavros Zenios, eds.: *The Performance of Financial Institutions: Efficiency, Inovation, and Regulation* (Cambridge University Press, Cambridge and New York).

Puri, Manju, 1994, The long term default performance of bank underwritten security issues, *Journal of Banking and Finance* 18, 397–418.

Puri, Manju, 1996, Commercial banks in investment banking: Conflict of interest or certification role? *Journal of Financial Economics* 40, 373–401.

Puri, Manju, 1999, Commercial banks as underwriters: Implications for the going public process, *Journal of Financial Economics* 54, 133–163.

Rajan, Raghuram, 2002, An investigation into the economics of extending bank powers, *Journal of Emerging Market Finance* 1, 125–156.

Rosenbaum, Paul, and Donald Rubin, 1983, The central role of the propensity score in observational studies for causal effects, *Biometrika* 70, 41–55.

Rubin, Donald, 1997, Estimating causal effects from large data sets using propensity scores, *Annals of Internal Medicine* 127, 757–763.

Santomero, Anthony M., and David L. Eckles, 2000, The determinants of success in the new financial services environment, *FRBNY Economic Policy Review*, October, 11–23.

Saunders, Anthony, and Ingo Walter, 1994, *Universal Banking in the United States. What Could We Gain? What Could We Lose?* (Oxford University Press, New York).

Smith, Randall, 2001a, Lucent deal shows Wall Street takes on greater risk, *Wall Street Journal* February 23, C1.

Smith, Randall, 2001b, Morgan Stanley injects about $2 billion into bank unit, aiming to boost lending, *Wall Street Journal* August 16, B7.

Stefanadis, Chris, 2004, Tying, exclusivity, and universal banking, Working paper, Rice University.

Yasuda, Ayako, 2005, Do bank relationships affect the firm's underwriter choice in the corporate-bond underwriting market? *Journal of Finance* 60, 1259–1292.

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1287 / JANUARY 2011

ARE UNIVERSAL BANKS BETTER UNDERWRITERS?

EVIDENCE FROM THE LAST DAYS OF THE GLASS-STEAGALL ACT

by Dario Focarelli,
David Marques-Ibanez
and Alberto Franco Pozzolo

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1287 / JANUARY 2011

ARE UNIVERSAL BANKS BETTER UNDERWRITERS?

EVIDENCE FROM THE LAST DAYS OF THE GLASS-STEAGALL ACT [1]

by Dario Focarelli [2], David Marques-Ibanez [3]
and Alberto Franco Pozzolo [4]

In 2011 all ECB publications feature a motif taken from the €100 banknote.

1 *Acknowledgements and disclaimer: Opinions expressed are those of the authors only and do not necessarily reflect the views or involve any responsibility of ANIA or of the European Central Bank. We are grateful to an anonymous referee as well as Patrick Bolton, Santiago Carbo, Bob DeYoung, Doug Evanoff, Craig Furfine, Luigi Guiso, Harri Huizinga, Paolo Mistrulli, Steven Ongena, Manju Puri, Richard Rosen and Lucy White for useful comments and/or discussions. We are in particular grateful to Barbara Sianesi for her advice on the matching procedure. We would also like to thank seminar participants at presentations and conferences at the Federal Reserve Bank of Chicago, at the University of Rome "Tor Vergata" Conference on Banking and Finance, at the EFMA Meeting, at the Global Finance Conference, at the XLIV Annual Conference of the Italian Economics Association for useful suggestions and at the joint BIS-JFI-CEPR conference on Systemic risk and financial regulation – causes and lessons from the financial crisis, Basel.*

2 *ANIA and Università di Roma "La Sapienza"; e-mail: dario.focarelli@ania.it*

3 *European Central Bank, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany; email: David.marques@ecb.europa.eu*

4 *Università del Molise, Centro d'Agliano and Mo.Fi.R.; e-mail: pozzolo@unimol.it*

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Internet
http://www.ecb.europa.eu

Fax
+49 69 1344 6000

Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website, http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html

ISSN 1725-2806 (online)

CONTENTS

Abstract: It has often been argued during the recent credit crisis that commercial banks' involvement in investment banking activities might have had an impact on the intensity of their underwriting standards. We turn to evidence from the period prior to the complete revocation of the Glass-Steagall Act in the United States and analyze whether investment banks or – section 20 subsidiaries of – commercial banks underwrote riskier securities. We compare actual defaults of these deals for an extensive sample of about 4,000 corporate debt securities underwritten during the period of the de facto softening of the Act's restrictions. Securities underwritten by commercial banks' subsidiaries have a higher probability of default than those underwritten by investment houses. This evidence is stronger in the case of ex-ante riskier and more competitive issues, and during the first years of bank securities' subsidiaries' entry into the market. Based on our results, it is not possible to reject that the repeal of the Glass-Steagall led to looser credit screening by broad (universal) banking companies trying to gain market share and/or to the lower initial ability of these banks to correctly evaluate default risk.

Keywords: Glass-Steagall Act, securities underwriting, default, investment banking.
JEL-classification: G21, G24, N22.

Non-technical summary

Over the last 30 years there has been a profound deregulation of investment banks' activities in the United States which lead to the creation of large financial institutions with a wide range of banking business. One of the most notorious examples of regulation setting up restrictions on banks' business models was the banking Act of 1933, commonly known as the Glass-Steagall Act, which imposed a functional separation between commercial and securities activities. At the time, the enactment of the Glass-Steagall was to a large extent motivated by concerns about conflicts of interest between the lending, underwriting and proprietary trading functions.

In the last decades of its existence the Glass-Steagall was progressively relaxed. In April 1987 the Fed allowed US commercial bank holding companies to establish affiliates which authorized to underwrite certain corporate securities. Finally, in 1999, the Glass-Steagall Act was formally repealed.

The dismantling of the Glass-Steagall Act was based on three main arguments. First, its revocation would allow banks to attain favorable economies of scope. Second, it would help banks to achieve greater opportunities for diversification derived from the different business lines with heterogeneous revenue cycles. Third, the repeal of the constraints introduced by the Glass-Steagall Act was expected to enhance the ability of US financial institutions to compete with foreign universal banks.

The rescission of the Glass-Steagall prohibitions was also buttressed by increasingly persuasive evidence from academic studies of the impact of broad banking on the pre-Glass-Steagall era. They consistently showed that securities underwritten by commercial banks' subsidiaries did not have a higher probability of default than those underwritten by investment banks.

We revisit the issue and present new empirical evidence on the impact of banks' business models on credit screening of corporate bonds. We examine empirically about 4,000 corporate bonds underwritten by investment houses and commercial banks' subsidiaries during the period of the progressive repeal of the Glass-Steagall (between 1991 and 1999) in the United States and analyze their ex-post credit quality using a unique data base of corporate defaults running until the end of 2008.

We find that over the ten years of the progressive repeal of the Act, debt securities issues underwritten by commercial banks had a higher probability of default than those underwritten by investment houses, the more so in the case of ex-ante riskier and more competitive issues, and during the first years of bank subsidiaries entry into the market. Based on our results, it is not possible to reject that the repeal of the Glass-Steagall led to looser credit screening by broad (universal) banking companies, although this seems to be due to banks' relative more aggressive underwriting standards in order to gain market share and/or the lower initial ability of these banks to correctly evaluate default risk rather than to conflicts of interest between the lending and underwriting functions. Our results do not question the elimination of the barriers between investment and commercial banking proposed by the Glass-Steagall Act. Rather, they point to the possibly perverse effects of allowing for increased competition without complementing it with more intensive scrutiny at the banking supervisory level.

I. Introduction

Over the last 30 years there has been a historical liberalization of banking markets in the United States. This liberalization included a “de facto” profound deregulation of investment banks’ activities which lead to the creation of large financial institutions with a wide range of banking business. While there have been a number of drivers of the financial crisis,[1] there has been a renewed interest – mostly from a financial stability perspective – on reassessing the usefulness of regulations imposing a functional separation between commercial and investment banking activities.

One of the most notorious examples of this regulation is the banking Act of 1933, commonly known as the Glass-Steagall Act, which abolished the securities affiliates of commercial banks (henceforth denoted as “banks” or “commercial banks” as opposed to “investment houses”) and imposed a separation or “firewall” between commercial and securities activities. At the time, the enactment of the Glass-Steagall was to a large extent motivated by concerns about the role of banks in the run up to the great depression and in particular about conflicts of interest between different banking activities and, in particular, between the lending, underwriting and proprietary trading functions.[2]

In the last decades of its existence the Act was progressively relaxed. In April 1987 the Fed allowed US bank holding companies to establish affiliates which authorized to underwrite certain corporate securities. Finally, in 1999, the Glass-Steagall Act was formally repealed with the passing and enactment of the Gramm-Leach-Bliley Act in 1999. At the time, three main arguments were put forward the repeal of the Glass-Steagall Act (Santos, 1998; Barth et al., 2000).

First, it would allow banks to attain favorable economies of scope. That is, for banks certain fixed costs of collecting, processing and assessing information or distributing financial services can be used across a range of financial services. These economies of scope were expected to enable financial holding companies to earn higher profits and pass along lower prices and to offer more product and service choices to their customers. In this respect, Kanatas and Qi (1997) analyze the trade-offs between the benefits derived from scope economies and the possible conflicts of interest and show that if commercial banks are allowed to underwrite securities, their economies of scope may enable them to gain all of the underwriting business of their loan customers by lowering credit standards.

Second, the dismantling of the Act was expected to allow banks to achieve lower levels of risk due to greater opportunities. These diversification opportunities for diversification

[1] They include, more prominently, financial innovation, the degree of competition in the banking system, external financial imbalances, the level of private sector debt, corporate governance in the banking sector, the relative tightness of monetary policy and the intensity of banking supervision among others (Acharya and Richardson, 2009).
[2] The separation of commercial and securities business was not a new idea. In fact it was the norm in many countries and had been American law and custom until the turn of the twentieth century, when securities affiliates of deposit taking institutions were allowed to operate (Mayer, 2009, and Gerschenkron, 1962).

derived from the different business lines with different revenues cycles were expected to translate in a lower profit variance for a broad banking company than for more specialized institutions (Cornett et al., 2002). Third, the repeal of the constraints introduced by the Glass-Steagall Act was expected to enhance US financial institutions' ability to compete with foreign universal banks.[3]

The repeal of the Glass-Steagall prohibitions was also supported by the increasingly persuasive evidence from academic studies of the impact of broad banking on the pre-Glass-Steagall era (Barth et al., 2000). Most empirical studies used data from the 1920s in the US (i.e. the pre-Glass-Steagall era) and showed that issues underwritten by the securities' subsidiaries of commercial banks did not have a higher probability of default than those underwritten by investment houses (Puri, 1994; Kroszner and Rajan, 1994).[4] In addition, securities underwritten by commercial bank subsidiaries at the time paid lower spreads, suggesting that investors recognized the stronger certification ability of commercial banks (Puri, 1996). Clearly, until very recently, the existence of the Glass-Steagall Act rendered it impossible to conduct a research on contemporary US firms using actual information on securities' defaults. This is the object of our research.

The role of banks' business models and their possible impact on the intensity of credit screening has become again a widely debated issue during the recent credit crisis. As a result, 10 years after the final repeal of the Glass-Steagall Act the separation between commercial banks and securities business is back into the political agenda.[5] A notorious, although very partial, reminiscent of the Glass-Steagall Act has been the so called "Volcker Rule", which was included in the Do Dodd–Frank Wall Street Reform and Consumer Protection Act signed into law by President Barack Obama on July 21, 2010.

We revisit the issue and present new empirical evidence on the impact of the functional separation of banks' business model on credit screening of corporate bonds. We examine 3,943 corporate bond issues underwritten by investment houses or commercial banks' subsidiaries during the period of the progressive repeal of the Glass-Steagall (between 1991 and 1999) and analyze their quality using a unique data base of corporate defaults running until the end of 2008. We find that debt security issues underwritten by commercial banks had a higher probability of default than those underwritten by investment houses.

Our results are in contrast with the so called "certification hypothesis", positing that commercial banks can benefit from informational economies of scope and therefore are in a

[3] This argument was often and empathically put forward politically (see Wilmarth, 2009).

[4] During the Nineties, a number of studies compared the characteristics of securities underwritten by commercial and by investment banks, respectively. A brief discussion is presented in section 2; a richer survey is in Crockett et al. (2003) and Morrison (2009).

[5] See for instance Senators' Cantwell and McCain legal initiative to reinstate the Glass-Steagall separation, which echoed proposals by senator Hinchey of New York as well as Exchange Commission ex-chairman Arthur Levitt ("An Odd Post-Crash Couple Spurning Obama, McCain and Cantwell propose resurrecting Glass-Steagall to break up Wall Street", Newsweek, 15th December, 2009).

better position than investment houses to correctly evaluate the securities that they underwrite.

We test two alternative explanations of our findings. First, based on the very rationale for the Glass-Steagall Act, we verify if they might be due to the presence of conflicts of interest derived from engaging simultaneously in lending and underwriting functions. Second, the pro-competitive effects of opening-up the security underwriting business and we analyze if commercial banks might have relatively lowered the underwriting standards and/or underestimated the securities' default risk, in their search for market share. We call the first the "conflicts of interest" hypothesis and the second the "risk underestimation" hypothesis.[6]

Our analysis points towards the second explanation. Securities issues underwritten by commercial banks have a higher probability of default particularly if they are ex-ante more opaque and riskier, if they are more competitive (i.e., if during the sample period the issuers used both banks and investment houses as underwriters; Kang and Liu, 2007), and if they are issued during the first part of our sample period. This is consistent with the hypothesis that the higher default rate was probably due to an underestimation of default risk on more opaque and contested issues, by part of new entrants into the market relatively lowering underwriting standards. In line with previous literature (Puri, 1996), we find instead very weak evidence consistent with the "conflict of interests" hypothesis. Our results do not question the elimination of the barriers between investment and commercial banking proposed by the Glass-Steagall Act. Rather, they point to the possibly perverse effects of allowing for increased competition without complementing it with more intensive scrutiny at the banking supervisory level.

The paper is organized as follows. Section 2 briefly describes the institutional environment in the U.S. market for corporate securities underwriting since the introduction of Section 20 subsidiaries. Section 3 briefly surveys the relevant literature. Section 4 describes the data used in the empirical analysis and presents some preliminary evidence based on the descriptive statistics. Sections 5 and 6 present the results of the econometric analysis. Section 7 concludes.

II. The institutional framework

The separation of commercial and investment banking (also called security business) activities has been the norm and custom in United States until the turn of the twentieth century. On the back of the stock market boom of the 1920s, commercial banks swarmed into the securities underwriting business. The McFadden Act in 1927 expressly allowed securities departments of national banks to underwrite securities (Peach, 1941). Prior to that, national banks were, in

---

[6] Conflicts of interest in the financial sector arise when a party to a transaction can gain by taking actions which are detrimental to its counterparty (see, Mehran and Stultz, 2007, and Crockett et al., 2003). The risk underestimation hypothesis is consistent, but not directly applicable, with theories suggesting that strategic interaction among asymmetrically informed banks may them to behave more aggressively and loose lending standards (Rajan, 1994, Ruckes, 2004, Dell'Ariccia and Marquez, 2006, and Gorton and He, 2008).

principle, barred from entering into the securities business. Yet in practice national banks were able to operate in the securities business by creating affiliates.[7] As a result, by 1929 these affiliates had attained half of the underwriting business in the United States.

The public outcry about the possible role of banks in the stock market crash of 1929 produced concerns about mixing investment and commercial banking activities.[8] This mood against the involvement of banks in securities business led to a Senate Banking and Currency Committee inquiry known as the Pecora committee, to ascertain conflicts of interest in the banking industry. The results showed indication of abuses by banks' securities affiliates and were cited by advocates of the Glass-Steagall Act, that was then passed in 1933. The Act banned banks from underwriting, holding or dealing in corporate securities, either directly or indirectly via securities affiliates. In particular, Section 20 of the Act ordered that: *"no member bank could be affiliated with any corporation, association or business trust engaged principally in the issue, flotation, underwriting, public sale, or distribution at wholesale or retail through syndicate participation of stocks, bonds, debentures, notes or other securities"*.[9]

Economically, one prime example of the possible conflicts of interest when marrying securities and commercial banking business is the possibility that universal banks might underwrite and push low-quality securities to investors (Morrison, 2009).[10] This would of course imply the existence of 'naive' investors – using the definition coined by Kroszner and Rajan (1994) – who would, in turn, buy these securities. Hence such a distinction between investment houses and banks implicitly suggests that they have different clienteles, who may in turn deserve different degrees of protection.[11] A problem with this interpretation is that it requires that investors are irrational and do not learn from past experience (Puri, 1999).

Another major argument supporting the functional separation provided by the Glass-Steagall Act was related to financial stability considerations. Namely, to keep deposit-taking institutions with access to the deposit insurance facilities out of activities that might lead to higher risk-taking and threaten the stability of the banking system (Freixas et al., 2007). The related possibility of facilitating "Too-Big-to-Fail (TBTF) Status" probably also supported the separation of business models as proposed by the Glass-Steagall Act. That is, the idea that the bankruptcy of certain very large and complex banks would be too harmful to the operation of the economy to be allowed to fail. This could exacerbate moral hazard problems possibly

[7] The Comptroller of the Currency ruled in 1902 that national banks were not allowed to engage in the securities business. Yet many banks were able to circumvent this rule by creating securities affiliates as state banks with their own capital, owned by the shareholders of the national bank (Morrison, 2009).
[8] See Willis and Chapman (1934) and Kennedy (1973).
[9] For a careful analysis of the institutional environment in the pre-Glass-Steagall period see Peach (1941), Carosso (1970), Kennedy (1971), Kelly (1985a), and Benston (1990).
[10] A contrary thesis is that the banks are better informed than investment houses on the financial conditions of their clients and are thus able to provide better certification of the securities that they underwrite.
[11] From this perspective the rationale for banning commercial banks from the underwriting business is in fact to be found in the need to protect their clients, who are presumed to be less able to evaluate the quality of the security issued.

creating incentives for excessive risk-taking (and/or regulatory capture) from those institutions (Mishkin, 2006).

The last argument in favor of the functional separation is behavioral in nature. It stresses the basic incompatibility of bringing commercial banking and investment banking activities together, as their incentives are not aligned, due to differences in the type of business conducted: investment banking is mostly based on fee-seeking brokerage activities in which short-term risk taking is paramount. On the other hand, commercial banking activities mostly hinge on maturity transformation of assets based on the long-term credit screening and monitoring of borrowers (Bhattacharya and Thakor, 1993). Combining both functions under the same institutional umbrella might shorten commercial bankers' incentives, lowering the intensity of screening and the long-term monitoring of borrowers.[12] This problem has probably become more acute in recent years as banks and markets have become increasingly integrated and co-dependent (Boot and Thakor, 2009).

Between 1933 and 1963 Glass-Steagall was fully enforced.[13] During this period, US investment banks became among the most competitive in the world and the share of financial intermediation grew rapidly as financial flows progressively shifted from the balance sheets of banks and other credit institutions to the financial markets. Starting in the mid-1960s, however, banks went back to the securities business, and eventually gained court authorization to underwrite municipal bonds. At the same time they were prevented from operating in most investment banking business.

In April 1987 the Fed allowed US commercial bank holding companies to establish affiliates (so-called "Section 20 subsidiaries") which were authorized to underwrite corporate securities. Two years later these affiliates were allowed to underwrite commercial paper and in 1990 permission was extended to equities. All these activities were allowed as long as they did not generate more than 5 per cent of the bank's total revenues (the ceiling was raised to 10 per cent in 1989 and 25 per cent in 1996). Finally, in 1999 the Gramm-Leach-Bliley Financial Services Modernization Act repealed the lingering legal barriers related to banks and investment houses' business separation.

III. The literature

The progressive loosening of the Glass-Steagall constraints generated a number of empirical studies analyzing the possible benefits of separating banks and investment houses' activities. These studies resorted to the analysis of the underwriting standards in the pre-Glass-Steagall era, and scrutinized two main hypotheses. The first one was that securities underwritten by

[12] Sheng (2009). For a more detailed analysis of the main arguments see Benston (1990).
[13] The Act impacted substantially on the state of banking sector and its future evolution. For example, in the aftermath of the set up, JP Morgan was split into the Morgan Bank and what later became Morgan Stanley.

banks might be riskier because of conflicts of interest between the lending and underwriting businesses. Alternatively, securities underwritten by banks could be safer, because commercial banks might be better informed on their creditors or simply have better credit assessment abilities than investment houses and thus can provide a more credible certification of credit risk.

The theoretical literature studying the implications of the presence of banks in the underwriting business is rather limited. Rajan (2002) argues that universal banks can be in conflict with their borrowers' interests because they can extract rents from their underwriting business using the monopoly power coming from their superior information on the borrowers and issuers' quality. Kanatas and Qi (1998) show that high quality issuers may have an incentive to be separated from low quality issuers. But since this can only be done using an investment house as the underwriter, they can choose to forgo the potential benefits of informational economies of scope with the lending services. The optimal choice of a high-quality issuer depends therefore on the trade-off between the conflicts of interests' costs and the benefits from informational economies of scope. In a later paper, Kanatas and Qi (2003) add one further dimension suggesting that universal banks have lower incentives to take their creditors directly to the markets and place their securities, and therefore may hamper their use of arms-length financing. Finally, Puri (1999) shows that the entry of commercial banks in the securities business can cause banks to lower the yields for underwritten securities particularly when information collection costs are high.

The empirical literature on the role of bank in the underwriting business is much richer. In a seminal paper, Kroszner and Rajan (1994) study the characteristics of a sample of industrial bonds underwritten by affiliates of commercial banks and investment houses in the first quarters of the years between 1921 and 1929. They obtain two major findings. First, the bonds originated by banks' affiliates were ex-ante safer (they had better ratings). This suggests that markets were probably aware of potential conflicts of interest and responded by imposing a 'lemons market' discount for banks on more information-intensive issues, forcing them to underwrite mostly the safer securities. Second, non-investment-grade bonds underwritten by banks' affiliates had fewer cumulative defaults in the period 1930-1940, both in number and in total value, and investment-bank underwritten bonds defaulted earlier in their life than affiliate-originated issues. This evidence tells against the thesis that commercial banks undertaking investment banking business were systematically defrauding their clients.

Puri (1994) and Ang and Richardson (1994) refine the analysis of Kroszner and Rajan (1994) and strengthen their results. Puri concentrates on the period subsequent to the McFadden Act of 1927, which explicitly allowed banks' affiliates to underwrite corporate securities, confirming that banks'-originated issues have a lower probability of default.[14] Using a wider set

[14] Kroszner and Rajan (1994) define bond issues as underwritten by commercial banks when any such institution is included in the syndicate. Puri (1994) includes only those where the affiliate is sole or lead underwriter.

of risk measures (ex-ante yield, default experience, ex-post market prices of bonds, stock prices of issuing companies), Ang and Richardson (1994) confirm issues underwritten by commercial banks did not performed worse than those underwritten by investment houses. Puri (1996) provides evidence of the commercial banks' greater certification abilities. Comparing two samples of securities underwritten in the pre-Glass-Steagall period, she finds that on average those originated by investment houses carry higher yields, confirming that commercial banks' greater certification ability outweighs the 'lemon' discount.

An alternative explanation of Puri's evidence is that banks have greater market power relative to their less sophisticated clientele so that they can place securities at higher prices (i.e. lower yields) than investment houses. But this is at odds with Kroszner and Rajan (1997) findings' suggesting that securities issued in the pre-Glass-Steagall period by the internal departments of commercial banks – for which the potential conflict of interest is even more severe – carry higher interest rates than those underwritten by the commercial banks' securities' affiliates. By confirming that markets consider an affiliate structure as an effective commitment mechanism, this result provides indirect evidence of the market's ability to discern – and price – the possibility of conflicts of interest.

Studying the Section-20 period, Roten and Mullineaux (2002) find no evidence of differences in the yield spreads between bonds underwritten by banks and investment houses between 1995 and 1998. However, Gande et al. (1997), show that securities underwritten by Section 20 subsidiaries of bank holding companies whose commercial bank affiliate is at the same time a lender to the bond issuer (i.e., those for which the potential for conflicts of interest is stronger) have lower interest rates, and more markedly so for lower-rated issuers. Moreover, commercial banks tend to underwrite smaller issues, offering further indirect confirmation of their greater ability to acquire and process information.[15] This result is partly confirmed by Saunders and Stover (2004), who find that when the same bank acts as underwriter and credit guarantor, interest rate spreads to the issuers are lower than average. Also in this direction, Drucker and Puri (2005) show that when an underwriter lends to an issuer around the time of an IPO (a practice known as 'tying'), the firm obtains a discounted interest rate on the loan, and the discount is greater for the more information-sensitive (i.e. non-investment-grade issues), which is consistent with the certification hypothesis.[16]

Evidence from different markets and countries is more mixed.[17] Hebb and Fraser (2002) show evidence of certification effects by commercial banks when studying the Canadian

[15] Gande et al. (1999) provide further evidence on the effects of the entry of Section 20 subsidiaries in the security underwriting market, showing that they have favored the reduction of underwriter spreads and ex-ante yields, and proportionally more so for lower-rated and smaller issues.
[16] A related issue is studied by Narayanan et al. (2004), who show that, in order to signal that they are not willing to exploit potential conflicts of interest to their advantage, banks acting as underwriters to their clients predominantly co-manage with a high reputation non-lending institution.
[17] A parallel strand of literature has studied the effects of the introduction mixing commercial and investment banking on bank risk and profitability (Morrison, 2009, and Drucker and Puri, 2007).

corporate bond market. Consistent with the US results, they show that yields on issues underwritten by Canadian commercial bank affiliates are lower than those on issues originated by independent investment banks. Also in this direction, Konishi (2002) finds no differences in the initial yields in bonds underwritten by commercial banks and investment houses in Japan, but finds lower default rates for the former.[18]

Evidence consistent with the prevalence of certification over conflicts of interest comes also from IPO underwritings. Schenone (2004) and Benzoni and Schenone (2010) show that IPOs underwritten by relationship banks are relatively less underpriced, confirming that they are not perceived as riskier than those underwritten by banks that have no lending relationships with the issuer. Similarly, studying a sample of over 2,000 seasoned equity issues in the US market between 1998 and 2006, Duarte-Silva (2010) finds that those underwritten by a lending-relationship bank had systematically higher announcement returns. In contrast, Ber et al. (2001), studying the Israeli market, find that issuing firms whose equity was underwritten by a commercial bank affiliate had worse stock market performance but better accounting profitability, which suggest the existence of both a certification effect from commercial banks as well as the existence of "naïve investors".

Finally, a recent paper by Kang and Liu (2007) takes a rather different perspective, showing that banks may have conflicts of interest with the issuers, rather than with the investors. Focusing on Japan between 1995 and 1997 – after Financial System Reform Act of 1993 allowed banks to provide investment banking services – they find evidence strongly supportive of the hypothesis that banks with stronger relationships with their clients used their market power in the lending business to force the issuers to pay higher yields on their securities, so as to attract a larger number of investor and gain market shares in the underwriting business, a result consistent with the theoretical framework of Rajan (2002).

We analyze the ex-post default rates of securities underwritten by banks and investment houses in the post-Glass-Steagall period. Namely in the following, we assess the impact of the repeal of the Glass-Steagall Act on underwriting standards by investment houses compared to banks undertaking investment banking business.

IV. Data and summary statistics

We analyze all nonconvertible fixed-rate corporate bonds issued in the United States between January 1, 1985, and December 31, 1999. Our data source is the Thomson Financial Database, a private commercial dataset on securities issues that provides, among other information, date of issuance, yield to maturity, credit rating, size, maturity and issuer's sector of economic activity,

[18] Results apply to Japan before the Second World War, when commercial and investment banking were not separated.

and is compiled from regulatory filings, news sources, company press releases, and prospectuses.

Information on defaults is from Moody's Corporate Default Risk Service Database, which includes global comprehensive information on defaults since 1970 for 11,000 corporate and sovereign entities. The data is derived from Moody's own proprietary database and is updated on a monthly basis. Our sample consists of defaults between January 1985 and December 2008. In total, 356 nonconvertible bonds in our sample defaulted. While some of them were issued by the same company, not all security issues by the same firm were underwritten by the same syndicate, justifying their inclusion in the analysis as distinct events. Missed interest or principal payments made up the majority of defaults. Less frequent reasons of default are filing for Chapter 11, filing for bankruptcy, distressed exchange and grace period default.

As opposed to underwritings from investment houses, we define bank underwritings as those by one of the subsidiaries (so called "Section 20" subsidiaries) of commercial banks (i.e. deposit-taking institutions with access to deposit insurance).[19]

For the empirical analysis we exclude the period before 1991, when the number of issues underwritten by banks was insignificant. In total, we have a sample of 3,943 fixed-rate corporate bonds issued between 1991 and 1999, 842 underwritten by Section 20 subsidiaries (21.4 per cent; 28.9 per cent in dollar terms) and 3,101 underwritten by investment houses. Banks' share in the underwritings' business changed significantly over the 1990's. In dollar terms, it rose from 18.2 per cent in 1991, to 42.9 per cent in 1995, and declined again to 35.4 per cent in 1999.

The default rate is significantly higher for bonds underwritten by banks than for investment houses, both in terms of dollar and number of issues. The average default rate for issues underwritten by banks is 11.9 per cent (10.3 in dollar terms), compared to 8.3 per cent for investment houses (7.6 percent in dollar terms, showing that defaulted issues underwritten by investment houses have a smaller size than their average issue). This difference is particularly striking for lower-grade bonds – i.e., those rated Baa2 and below – while it has the opposite sign for higher-grade securities. Of the 1,995 lower-grade issues in our sample (Table 1), 1,522 were underwritten by investment houses and 473 by commercial banks (Table 2). The rate of default of issues underwritten by investment houses is 12.8 percent, much lower than the 18.2 per cent of those underwritten by banks.

[19] The Federal Reserve Board granted approval to a bank holding company to engage through a so-called Section 20 subsidiary in underwriting and dealing in securities that a member bank may not underwrite or deal in directly. Such a subsidiary is called a Section 20 subsidiary in reference to a section of the Glass-Steagall Act that limited affiliations between certain securities companies and member banks. The list of Section 20 subsidiaries is from Cornett et al. (2002).

Another striking difference is between the performance of bonds issued in the first and in the second part of our sample period, possibly due to the different macroeconomic conditions. Securities issued between 1991 and 1994 show a rate of default of 11.4 per cent, while those underwritten between 1995 and 1999 only have a rate of default of 7.3. Securities underwritten by banks had a much higher rate of default than those underwritten by investment houses (Table 1). The fact that the differences in the ex-post performance of securities issued by banks and investment houses is larger for lower-grade securities is consistent with both the conflict of interest and the "risk underestimation hypothesis". That is, due to increased competition commercial banks lowered underwriting standards to attain market share or simply due to their lower credit assessment abilities, looser standards were "ex-ante" underestimated by investors. However, the fact that the difference is lower in the second part of our sample period, when banks had already gained market shares in the underwriting business and were likely to be more cautious due to reputational issues, is more supportive of the risk underestimation hypothesis.[20]

Table 2 provides some additional evidence consistent with the risk underestimation hypothesis, showing that the difference in the default rates of securities underwritten by banks and investment houses is much starker for the sub-sample of competitive issues (i.e., those made by issuers which used both banks and investment houses as underwriters during our sample period; Kang and Liu, 2007).[21] If the reason for the higher default rate of bank underwritten issues were conflicts of interest, we would expect that banks push especially the securities issued by their longer term borrowers, for which they have most likely obtained through time a reliable set of soft information, but that are less likely to switch underwriter from banks to investment houses. In the next Section we present the results of a more formal analysis of the default probability of securities underwritten by investment houses and commercial banks.

V. Regression results

V.a Baseline specification

Our basic specification brings back the empirical methods used to analyze the pre-Glass-Steagall era. Hence we start by building on the model by Kroszner and Rajan (1994) and test whether securities underwritten by commercial banks' subsidiaries have a higher ex-post probability of default than those underwritten by investment houses. We start by estimating the following binary choice model, using a probit specification:

[20] Our data set includes relatively few instances of securities underwritten by banks that were at the same time lending to the issuers on the syndicated loan market. When excluding these issues, the patterns in the default rates of securities underwritten by banks and investment houses remain the same described above, proving that conflicts of interest cannot be taken as their unique explanation.
[21] Earlier literature shows that borrowers issuing securities tend often to hire the same investment bank often demanding exclusivity in the bond market (Yasuda (2005)) and to lesser extent in the equity market (Ljungqvist, Marston, and Wilhelm (2006))

$$Pr\ (Y_{i,j,t} = k) = f\ (X_{i,t}, D_j, Z_i, T_t),\ k = 0, 1; \tag{1}$$

where: $Y_{i,j,t} = 1$ if security issue i underwritten by bank j at time t defaulted before maturity and $Y_{i,j,t} = 0$ otherwise; $X_{i,t}$ are characteristics of the issue i at time t (i.e., size, rating, maturity, gross spread over benchmark, issuer's sector of economic activity); D_j is a dummy taking the value of 1 if there is at least one Section 20 subsidiary among the banks leading the underwriting syndicate; Z_i is a set of dummies for the sector of economic activity of the issuer; T_t are year dummies. Unlike Kroszner and Rajan (1994), we also include information on the issue's gross spread, to account for a higher probability of default known to investors at the time of the issue but not included on the rating and other issue characteristics.[22] All standard errors are calculated using the procedure of White (1980) to correct for heteroskedasticity and clustering at the issuer level to account for the fact that the behaviors of bonds issued by the same borrower are not independent.

Panel A of Table 3 reports the results of the estimation of the basic specification for equation (1). The regression includes time and industry dummies, as well as the issue's size, gross spread with respect to the benchmark, maturity and rating. Estimates are conducted on a sample of 3,279 security issues, from 1991 to 1999.[23]

The coefficient of the dummy variable for issues underwritten by commercial banks is consistent with the hypothesis that securities issues underwritten by commercial banks have a higher probability of default. That is, on average, issues underwritten by commercial banks have a higher probability of underwriting a security that eventually went into default. This result reinforces what was already apparent from the descriptive statistics, since the higher default probability is confirmed also after controlling for the characteristics of the deal, issuer and macroeconomic characteristics. The additional controls show that the size of the issue has no impact on its default probability. On the contrary, securities with higher yields have a higher default probability, even after controlling for rating and duration.

In Panel B we have distinguished between high-grade (i.e. low level of credit risk) and low-grade (i.e. higher level of credit risk) securities, including an interaction term. High-grade ratings are those classified by Moody's between Aaa and Baa1. Low-grade ratings are those classified by Moody's as Baa2 and below. The coefficient of the interaction term between the dummy for banks and that for lower-grade securities is positive, significantly different from zero, and relatively large. On the contrary, the coefficient of the interaction term between the dummy for banks and that for higher-grade securities is negative, but it is not statistically

[22] Identical results are obtained excluding the gross return from the set of the explanatory variables.
[23] Of the total number of issues in the period (3,943; Table 1), 308 observations are dropped because we do not have any defaults for issues with a rating Aaa, Aa1 and Aa2, Ca, Caa2 and not rated, while 337 are dropped because we do not have defaults for firms with one-digit SIC code 6, i.e. firms operating in regulated industries.

significant. The main results are therefore driven by differences in credit screening abilities on the underwriting of corporate bonds with relatively high levels of credit risk (i.e. lower ratings).

V.b Alternative econometric specifications

In addition to the binary choice model presented above, we checked the robustness of our results using two alternative econometric specifications.

First, in order to account more carefully for the different duration of the bonds linked to the credit worthiness of each bond, we estimate a survival model by the method of proportional hazards regression first proposed by Cox (1972):

$$\lambda(t_i) = e^{-(\beta' x_i + \gamma D_i + \delta Z_i + \xi T_t)} \lambda_0(t_i) \qquad (2)$$

where: λ_0 is the 'baseline' hazard and all other explanatory variables are as indicated above. For defaulted bonds, we define the duration as the period from the date of issuance to the date of default; for non-defaulted bonds, as the period from the date of issuance to the date of repayment. As before, the specification includes, as deal control variables, issue's maturity, rating, size and gross spread with respect to the benchmark as well as time and issuer's industry dummies.

The results for the coefficient of the dummy for securities underwritten by banks reported in Panels C and D of Table 3 confirm the results obtained with the binary choice model. Similarly, they also confirm that the overall impact comes from the performance of low-rated securities (Panel D). The impact of the control variables are also unchanged.

Second, to correct for the possible bias induced by the use of a parametric specification of the probability of default linked to differences in the composition of the selected populations of deals according to the type of underwriter biasing our results, we adopt a propensity score matching method (Rubin, 1979).[24] In practice, we split our sample between bonds underwritten by banks ('treated' observations) and bonds underwritten by investment houses ('untreated' or 'control' observations), match each 'treated' observation with a set of 'untreated' observations chosen so as to be as similar as possible to the 'untreated' ones and then compare the probability of default between the two groups. The matching is based on the propensity score estimated from a first stage binary choice model using size, rating, maturity, gross spread over the benchmark and issuer's sector of economic activity as explanatory variables. The nearest neighbors approach is used.[25]

---

[24] These methods, first introduced in the medical sciences, are now becoming increasingly popular also in economics. They lend themselves naturally to our analysis, because they focus precisely on non-random selection. For a recent survey, see Blundell and Costa Dias (2002). The routine we used for estimations is PSMATCH2, a Stata module by Leuven and Sianesi (2003). For an early application to finance see Villalonga (2004).

[25] We found similar results using the kernel weights method suggested by Heckman et al. (1997).

Panel A of Table 4 shows that the probability of default for securities underwritten by Section 20 subsidiaries (the 'average treatment on the treated' in the column labeled 'treated') is larger than for securities underwritten by investment banks (the 'average treatment on the treated' in the column labeled 'controls') and the difference is statistically significant.[26] Even stronger results are obtained for low-grade (i.e. with higher credit risk) securities, but the difference between the two effects is not statistically significant (panel B). Finally, for higher-grade (i.e. low credit risk) securities (panel C), the average treatment effect for 'treated' observations is equal to 0.04, which is almost double than for the control sample (0.023), and it is statistically significant.

V.c Discussion

In summary, the results obtained with all three statistical methodologies presented above lend strong support to the view that securities underwritten by the commercial bank securities subsidiaries have a higher probability of default than those underwritten by investment houses. This is in contrast with the certification hypothesis, according to which commercial banks are able to screen credit quality relatively better than other institutions before passing the underwritten securities through to investors.

Our results are consistent with findings of Kang and Liu (2007), who do not find a certification view of bank underwriters. Instead they show that banks that are more rapidly expanding their underwriting business face a greater need to provide investors with high-reward investment opportunities and discount the prices of bonds in order to attract potential investors

They are also related to those by Ber et al. (2001), who suggest that that fund managers are able to find and to exploit "naïve" or "unsuspecting" retail investors who are not able to distinguish and fully price ex-ante the impact of institutional underwriting (i.e. whether the securities were underwritten by banks or investment houses) on ex-post credit quality.

However, the baseline results presented above provide no clear explanation of the reasons why securities underwritten by banks are ex-post more likely to default than those underwritten by investment houses. In particular, it is not clear whether this depends on the presence of conflicts of interests or on other reasons leading banks to underestimate credit risk. In the following, we discuss some possible alternative explanations of our results, by delving further into the characteristics of bonds issued by banks and investment houses.

[26] Standard errors are computed by bootstrapping with 200 repetitions and significance by using the bias-corrected confidence interval.

VI. Additional evidence

VI.a Ex-post defaults

As it was already apparent from the descriptive statistics, the difference between the probabilities of default of securities underwritten by banks and investment houses is more striking in the first part of our sample period, a result leading support to the risk underestimation hypothesis.

In Table 5 we split our sample between issues underwritten between 1991 and 1994 and those underwritten between 1995 and 1999. Consistent with the evidence reported in the descriptive statistics, the coefficient of the dummy for issues underwritten by banks stronger when estimated in the first part of our sample (Panel A), but it becomes statistically insignificant when estimated in the second part of the sample (Panel B). Further, Panel D and E confirm that also in this case the results are driven by the differences in credit screening abilities on the underwriting of corporate bonds with relatively high level of credit risk (i.e. lower ratings). These results suggest that, initially, the banks that were entering the underwriting business underestimated the credit risk of the more opaque securities (i.e., those with lower credit ratings). Since we control for rating in our regressions, risk underestimation took place within each rating bucket, and for given size and maturity. Ex-post, risk underestimation led to the higher default incidence of securities underwritten by banks. In the second part of our sample period, when banks had already gained a significant share of the market and reputational issues might have become more compelling, relative risk underestimation nearly disappeared.

Evidence obtained from splitting the sample in two periods might also consistent with the conflicts of interest hypothesis, if one assumes that banks initially underwrote securities for the borrowers that they knew best, for which the potential for conflicts of interest was stronger, and only later offered their services to a wider clientele.

To contrast this explanation we have further focused on a sub-sample of competitive issues, defined as those made by issuers that used both banks and investment houses as underwriters during our sample period (Kang and Liu, 2007). On one hand, it is more likely that banks take an aggressive strategy to gain new clients from investment houses within the group of more competitive issues, possibly requiring lower yields and therefore underpricing risk. On the other hand, it is less likely that banks have strong informational advantages with respect to investment houses on these issuers, since these clients are less reliant on strong relationships with a given financial intermediary. Finding that competitive issues underwritten by banks have a higher default incidence than those underwritten by investment house would therefore provide additional support to the risk underestimation hypothesis, in contrast to the conflict of interest hypothesis.

Table 6 shows that this is indeed the case. The coefficient of the dummy for bank underwriting in the sub-sample of competitive issues is positive and strongly significant (Panel

A), which is not the case for uncompetitive issues (Panel B). Panels D and E confirm once again that the general results are driven by the differences in credit screening standards or abilities on the underwriting of corporate bonds with relatively high level of credit risk (i.e. lower ratings).

Table 7 provides some additional results consistent with the risk underestimation and competition on "de novo" entrants hypotheses, showing that foreign banks, that are more likely to incur in risk underestimation or in a looser underwriting standards to gain market shares, have a much higher probability of underwriting securities that ex-post defaulted (Panel A).

Panel B looks instead at the impact of maturity at issuance on debt defaults and bank underwritings. In this respect, the literature emphasizes that under information asymmetry, credit quality of the issuer influences debt maturity (see Flannery, 1986, and Berger et al., 2005). Firms with lower credit quality are more likely to be forced by the market to issue short-term debt, to minimize moral hazard problems, while firms with higher credit quality would tend to issue long-term debt. Consistent this prediction, we find that the dummy for short term (below 5 years) securities underwritten by banks is significant, while that for longer term securities (above 10 years) was statistically insignificant. Also in this case, the evidence seems in contrast with the conflicts of interest hypothesis between underwriters and investors, since banks would be more willing to push longer-term low-quality securities to investors, gaining more time to recover their risky lending expositions.

VI.b Ex-ante conditions

In a seminal contribution, Gande et al. (1997) found that interest rates on securities underwritten by banks were significantly lower than those on securities underwritten by investment houses, a result providing strong evidence in favor of the certification hypothesis. While our results on ex-post defaults are clearly at odds with such hypothesis, it might still be the case that investors believed in banks' superior certification ability, and therefore accepted lower returns on the securities that they underwrote. Banks might have then exploited the investors' belief of their superior certification ability, and required lower interest rates on the issues that they underwrote, in order to attract new issuers.

We have tested this hypothesis verifying if the gross spread on securities underwritten by banks is significantly different from that on securities underwritten by investment houses, controlling for all the characteristics of the issue considered in the previous regressions. Table 8 shows that this is not the case (Panel A), neither distinguishing the effect on high-grade and low-grade securities (Panel B).[27] Interestingly, securities underwritten by foreign banks are instead required to pay a premium to investors (Panel C), although we already know from our previous analysis on the probability of default that this is insufficient to compensate for their

[27] Roten and Mullineaux (2002) find similar results analyzing a different sample than Gande et al. (1997).

higher riskiness, because in those regressions we did control also for the gross yield on the bond, and still estimated a positive and significant coefficient for the dummy of foreign banks.

Next, we ran a counterfactual exercise verifying if ex-ante spreads charged on those securities that eventually defaulted were higher than average. Effectively this would detect whether the market had some information on the credit risk which was included in the spreads but is not captured by our controls. Interesting, we find that this is indeed the case, as shown by the positive and significant coefficient of the dummy for issues that eventually defaulted (Panel D). However, this effect is common to all issues underwritten by both banks and investment houses, as shown by the insignificant coefficient of the term accounting for the set of defaulted issues underwritten by a Section 20 subsidiary.

VI.c Further controls

In a number of unreported regressions, available from the authors upon request, we have also checked some additional implications of the conflicts of interest and the risk underestimation hypotheses. First, we matched our information with those on syndicated lending obtained from Loanware, a commercial database provided by Dealogic, and we have singled the cases in which banks had an ongoing lending relationship at the time of the issue or had had one in the previous years but we do not find evidence suggesting that earlier lending relationships had an impact on our results.

Second, we verified if issues underwritten by banks were ex-ante riskier than those underwritten by investment houses, estimating a binary choice model for the probability that a security is underwritten by a bank. A possible explanation for the higher default rate on issues underwritten by banks – particularly for lower quality deals – is that the ratings assigned to those securities were systematically different from those of bonds underwritten by investment houses. This might happen for two different reasons, both assuming that commercial banks are unable to gather all the information necessary to fully discriminate between different securities. On the one hand, rating agencies might assume that commercial banks have an incentive to lower their screening incentives when underwriting securities. For instance, they might assume that there is an ex-ante conflict of interest and that some banks might misrepresent the issues' quality to sell it to unsuspecting investors. In this case, they would tend to assign de facto an inferior rating to bonds originated by commercial banks. In other words, rating agencies would apply a 'lemon' discount to issues underwritten by banks (Kroszner and Rajan, 1994). Alternatively, rating agencies might assign better ratings to securities underwritten by Section 20 subsidiaries of commercial banks because they believe that, on average, commercial banks are simply better at screening credit risk than investment houses, because they have private information on the issuer, as suggested by the certification effect.

Clearly a 'lemon' discount on bank-originated issues would not imply a bias in our results.[28] This, however, would not hold if the rating assignment process assumed the existence of a certification effect on the part of commercial banks. Under this second hypothesis, other things being equal, bank-originated securities would then be assigned better ratings than those originated by investment houses. Lower-grade issues underwritten by commercial banks would then be of relatively lower quality and, as such, would have a higher default rate. However, we found evidence against this hypothesis, since all the dummies rating buckets in the model estimating the probability that an issue is underwritten by a bank or by an investment house are statistically insignificant, suggesting that banks are not more likely to underwrite riskier securities than investment houses.

Third, following Gande et al. (1997), we used the residuals from the binary choice model on the probability that a security is underwritten by a commercial bank as a proxy for underwriters' private information, and we plugged them into our baseline specification for the probability of default. Their effect is statistically insignificant, suggesting that neither banks nor security houses made intentional use of private information, neither for certification purposes nor to exploit potential conflicts of interest.

Fourth, we verified if the progressive entry of banks in the security business affected the default rates. Since in all our previous specifications we included year dummies among the explanatory variables, we replaced them with the aggregate market share of banks in security underwriting. Our results remained unchanged.

Fifth, we checked whether the interest rates on bank underwritten securities were significantly lower from the averages for the subsamples of competitive issues, lower-grade issues, and issues underwritten before 1995, as it might have been the case if banks had purposely tried to use their market power to place bonds at lower prices, to gain market shares among issuers. In none of these cases the difference was statistically significant.

VII. Conclusions

Our results show that securities underwritten by banks have a significantly higher default rate than those underwritten by investment houses, and that this cannot be explained by ex-ante publicly available information on the characteristics of the issues or the issuers. This result is in contrast with the certification hypothesis, stating that banks benefit from informational economies of scope and therefore can assess more precisely the riskiness of the securities that they underwrite.

[28] In presence of a 'lemons market' discount, affiliate-originated issues would be systematically rated lower than issues underwritten by commercial banks and therefore, within each rating class, they should have on average a lower probability of default (see Kroszner and Rajan, 1994, pp. 822-823).

We have proposed two alternative explanations of our findings. First, following the very rationale for the Glass-Steagall Act, we posited that banks might exploit their private information and misrepresent the issues' quality to sell it to unsuspecting investors, possibly using the proceedings to pay back their own loans. However, we found no convincing evidence consistent with this hypothesis.

Second, we suggested that banks might have been induced to underestimated default risk in their search for market shares. We found evidence consistent with this hypothesis, especially for securities that were low-graded, issued by corporations that more aggressively searched for better conditions, and in the first years since the admission of banks into the security business. In other words, banks had to be initially more aggressive than investment banks houses in order to gain market share, and in pursuing this objective they might have loosened their credit standards excessively.

The recent crisis has revived the interest on the possible impact of the repeal of the Glass-Steagall Act on banks' incentives. Our results do not question the elimination of the barriers between investment and commercial banking proposed by the Glass-Steagall Act. Rather, they point to the possibly perverse effects of allowing for increased competition without complementing it with more intensive scrutiny at the supervisory level.

Partly due to de-regulation and partly due to financial innovation, the traditional bank lending and securities business have become increasingly intertwined (Thakor and Boot, 2009) and a number of investment banks have become systemic in nature. Hence further empirical evidence on the competitive dynamics in the investment banking business and its possible impact on risk-taking incentives is, to our mind, warranted. In particular, it would be interesting to know whether a more crowded market for investment bank business – including universal banks with access to deposits and central banking facilities – might impact banks' risk taking incentives. At the same time, it would be relevant to understand how commercial banks further involvement in securities business might be affecting their broad credit screening intensity in traditional lending activities.

References

Acharya, V. and M. Richardson (2009), "Causes of the Financial Crisis", *Critical Review*, 1, 195-210.

Ang, J.S. and T. Richardson (1994), "The Underwriting Experience of Commercial Bank Affiliates Prior to the Glass-Steagall Act", *Journal of Banking and Finance*, 18, 351-395.

Barth, J. R., R.D. Brumbaugh, and J.A. Wilcox (2000), "The Repeal of Glass-Steagall and the Advent of Broad Banking", *Journal of Economic Perspectives*, 14, 191–204.

Benston, G.J. (1990), *The Separation of Commercial and Investment Banking: The Glass-Steagall Act Reconsidered*, New York, Oxford University Press.

Benzoni, L. and C. Schenone (2010), "Conflict of Interest and Certification in the US IPO market", *Journal of Financial Intermediation*, 19, 235-254.

Ber, H., Y. Yafeha, and O. Yosha (2001), "Conflict of Interest in Universal Banking: Bank Lending, Stock Underwriting, and Fund Management", *Journal of Monetary Economics*, 47, 189-218.

Berger, A.N., M.A. Espinosa-Vega, S.W. Frame, and N.H. Miller (2005), "Debt Maturity, Risk, and Asymmetric Information", *Journal of Finance*, 60(6), 2895-2923.

Bhattacharya, S., and A.V. Thakor (1993), "Contemporary Banking Theory", *Journal of Financial Intermediation*, 3-1, 2-50.

Blundell, R., and M. Costa Dias (2002), "Alternative Approaches to Evaluation in Empirical Microeconomics", Cemmap working paper 10/02.

Boot, A.W.A. and A. V. Thakor (2009), "The Accelerating Integration of Banks and Markets and its Implications for Regulation" in A. Berger, P. Molyneux and J.Wilson editors, Handbook of Banking, Oxford University Press.

Carosso, V.P. (1970), *Investment Banking in America: A History*, Cambridge, Massachussets, Harvard University Press.

Cornett, M., E. Ors, and H. Tehranian (2002), "Bank Performance around the Introduction of a Section 20 Subsidiary", *Journal of Finance*, 57, 501-521.

Cox, D. (1972), "Regression Models and Life Tables", *Journal of the Royal Statistical Society, Series B*, 34, 187-220.

Crockett, A., T. Harris, F. S. Mishkin, and E. N. White (2003), *Conflicts of Interest in the Financial Services Industry: What Should We Do About Them?*, Geneva Reports on the World Economy, 5.

Dell'Ariccia, G. and R. Marquez (2006), "Lending Booms and Lending Standards", *Journal of Finance,* 61, 2511-2546.

Drucker, S., and M. Puri (2005), "On the Benefits of Concurrent Lending and Underwriting", *Journal of Finance,* 60, 2763-2799.

Drucker, S. and M. Puri (2007), "Banks in Capital Markets: A Survey" In: Eckbo, B.E. Editor, *Handbook in Corporate Finance: Empirical Corporate Finance,* Elsevier.

Duarte-Silva, T. (2010), "The Market for Certification by External Parties: Evidence from underwriting and banking relationships", *Journal of Financial Economics*, 98, 568-582.

Fang, L. (2005), "Investment Bank Reputation and the Price and Quality of Underwriting Services," *Journal of Finance*, 60, 2729-2761.

Flannery, M.J. (1986), "Asymmetric Information and Risky Debt Maturity Choice", *Journal of Finance*, 41, 19-37.

Freixas, X., G. Lóránth, and A.D. Morrison (2007), "Regulating Financial Conglomerates", *Journal of Financial Intermediation*, 16, 479–514.

Gande, A., M. Puri, A. Saunders, and I. Walter (1997), "Bank Underwriting of Debt Securities: Modern Evidence", *Review of Financial Studies*, 54, 1175-1202.

Gande, A., M. Puri, and A. Saunders (1999), "Bank Entry, Competition, and the Market for Corporate Securities Underwriting", *Journal of Financial Economics*, 54, 165-195.

Gerschenkron, A. (1962), *Economic Backwardness in Historical Perspective*, Cambridge, Mass: Harvard University Press.

Gompers, P.A. and J. Lerner (1999), "Conflict of Interest in the Issuance of Public Securities: Evidence from Venture Capital", *Journal of Law and Economics*, 42, 1-28.

Gorton, G. and P. He (2008), "Bank Credit Cycles", *Review of Economic Studies*, 75, 1181-1214.

Hebb, G.M. and D.R. Fraser (2002), "Conflict of Interest in Commercial Bank Security Underwritings: Canadian Evidence", *Journal of Banking and Finance*, 26, 1935-1949.

Heckman, J., H. Ichimura, and P. Todd (1997), "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, 64, 605-654.

Kanatas, G. and J. Qi (1998), "Underwriting by Commercial Banks: incentive Conflicts, Scope Economies, and Project Quality", *Journal of Money, Credit, and Banking*, 30, 119-133.

Kang. J-K. and W-L. Liu (2007), "Is Universal Banking Justified? Evidence from bank underwriting of corporate bonds in Japan", *Journal of Financial Economics*, 84, 142-186.

Kelly, E.J. (1985), "Legislative History of the Glass-Steagall Act", in I. Walter (ed.), Deregulating the Wall Street: Commercial Bank Penetration of the Corporate Securities Market, John Wiley and Sons.

Kennedy, S.E. (1973), *The Banking Crisis of 1933*, University of Kentucky Press.

Konishi, M. (2002), "Bond Underwriting by Banks and Conflicts of Interest: Evidence from Japan During the Pre-War Period", *Journal of Banking and Finance*, 26, 767-793.

Kroszner, R.S. and R.G. Rajan (1994), "Is the Glass Steagall Act Justified? A Study of the US Experience with Universal Banking Before 1933", *American Economic Review*, 84, 810-832.

Kroszner, R.S. and R.G. Rajan, (1997), "Organization Structure and Credibility: Evidence from Commercial Bank Securities Activities before the Glass-Steagall Act", *Journal of Monetary Economics*, 39, 475-516.

Leuven, E. and B. Sianesi, (2003), "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing", http://ideas.repec.org/c/boc/bocode/s432001.html.

Ljungqvist, A., F. Marston, and W.J. Wilhelm (2006), "Competing for Securities Underwriting Mandates", *Journal of Finance*, 61, 301-340.

Mayer, M. (2009), "Glass-Steagall in Our Future: How Straight, How Narrow", Network Financial Institute Policy Brief, University of Indiana 2009-07.

Mehran, H. and R. Stulz (2007), "The Economics of Conflicts of Interest in Financial Institutions", *Journal of Financial Economics*, 85, 267-296.

Mishkin, F. S. (2006), "How Big a Problem is Too Big to Fail?", A Review of Gary Stern and Ron Feldman's *Too Big to Fail: The Hazards of Bank Bailouts*, *Journal of Economic Literature*, 91, 98-1004.

Morgan, D. (2002), "Rating Banks: Risk and Uncertainty in an Opaque Industry", *American Economic Review*, 92, 874-888.

Morrison, A.D. (2009), "Universal Banking", in A. Berger, P. Molyneux and J.Wilson editors, Handbook of Banking, Oxford University Press.

Morrison, A.D. and W.J. Wilhelm (2007), *Investment Banking: Institutions, Politics and Law*, Oxford University Press, Oxford, UK

Narayanan, R.P., K.P. Rangan, and N.K. Rangan (2004) "The Role of Syndicate Structure in Bank Underwriting", *Journal of Financial Economics*, 72, 555-580.

Peach, W.N. (1941), *The Security Affiliates of National Banks*, Baltimore, The Johns Hopkins University Press.

Puri, M. (1994), "The Long Term Default Performance of Bank Underwritten Security Issues", *Journal of Banking and Finance*, 18, 398-418.

Puri, M. (1996), "Universal Banks in Investment Banking: Conflict of Interest or Certification Role?", *Journal of Financial Economics*, 40, 373-401.

Puri, M. (1999), "Universal Banks as Underwriters: Implications for the Going Public Process", *Journal of Financial Economics*, 54, 133-163.

Rajan, R. (1994), "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence", *Quarterly Journal of Economics*, 109, 399-441.

Ruckes, M. (2004), "Bank Competition and Credit Standards", *Review of Financial Studies*, 17, 1073-1102.

Roten, I.C. and D.J. Mullineaux (2002), "Debt Underwriting by Commercial Bank-Affiliated Firms and Investment Banks: More Evidence", *Journal of Banking and Finance*, 26, 689-718.

Rubin, D.B. (1979), "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies", *Journal of the American Statistical Association*, 74, 3, 8-329.

Santos, J.A.C. (1998), "Commercial Banks in the Securities Business: A Review", *Journal of Financial Services Research*, 14, 35-60.

Saunders, A. and R.D. Stover (2004), "Commercial Bank Underwriting of Credit-Enhanced Bonds: are there Certification Benefits to the Issuer?", *Journal of International Money and Finance*, 23, 367-384.

Schenone, C. (2004), "The Effect of Banking Relationships on the Firm's IPO Underpricing", *Journal of Finance*, 59, 2903-2958.

Sheng, A. (2009), *From Asian to Global Crisis*, Cambridge University Press, 2009, 326.

Villalonga, B. (2004), "Does Diversification Cause the 'Diversification Discount?", *Financial Management*, 33, 5-27.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, 817-838.

Willis, H.P. and J.M. Chapman (1934), *The Banking Situation,* Columbia University Press.

Wilmarth, A.E. (2009), "The Dark Side of Universal Banking: Financial Conglomerates and the Origins of the Subprime Financial Crisis", *Connecticut Law Review*, 41, 4.

Yasuda, A. (2005), "Do Bank Relationships Affect the Firm's Underwriter Choice in the Corporate Bond Underwriting Market?, *Journal of Finance*, 60, 1259-1292.

Table 1

Number and volume of securities issued by year, by type of underwriter and by default

The table gives the number and the total aggregate values of the securities in our sample. They are issued between January 1, 1991 and December 31, 1999, with maturity longer than December 31, 1999. It also includes the default rates in percentage of those securities according to whether they were underwritten by investment houses or commercial banks. Commercial bank underwritings are those by one of the Section 20 subsidiaries. Data on security issues are from the Thomson Financial Database, defaults are from Moody's Default Report. High-grade ratings are those classified by Moody's between Aaa and Baa1. Low-grade ratings are those classified by Moody's as Baa2 and below.

	Total market		**Market share of bank underwritings** *(in percentage)*		**Default rate** *(in percentage)*			
					Investment houses		**Banks**	
	Number of issues	Amount *(mill. $)*	Number of issues	Amount	Number of issues	Amount *(mill. $)*	Number of issues	Amount *(mill. $)*
			Panel A: Total securities issued					
1985	81	11,460			32.1	26.3		
1986	289	43,633			17.6	15.6		
1987	195	50,122			23.1	31.5		
1988	150	37,977			14.7	12.6		
1989	139	41,146	2.2	0.8	15.4	10.3		
1990	104	21,061	1.9	4.4	3.9	4.2		
1991	259	57,740	13.5	18.2	4.5	4.2	8.6	11.3
1992	464	106,011	23.9	31.1	12.2	10.5	14.4	9.6
1993	624	156,595	25.3	29.2	11.2	9.4	14.6	14.7
1994	301	64,763	31.2	40.5	11.1	9.7	19.1	24.5
1995	329	84,394	42.9	49.5	8.0	7.8	7.1	4.3
1996	411	113,820	23.8	25.4	7.3	5.7	13.3	17.7
1997	321	115,048	16.5	22.0	4.9	3.7	9.4	5.0
1998	630	202,796	7.8	15.0	5.7	3.0	6.1	9.8
1999	604	275,378	17.1	35.4	8.8	12.8	8.7	6.4
91-99	3,943	1,176,544	21.4	28.9	8.3	7.6	11.9	10.3
85-90	958	205,400	0.5	0.6	17.7	17.4		
91-94	1,648	385,109	24.2	30.0	10.2	8.8	15.1	15.1
95-99	2,295	791,435	19.3	28.3	6.9	7.0	9.0	7.8
			Panel B: High grade securities					
85-90	481	86,522	0.8	1.3	4.4	4.1		
91-94	828	193,209	19.2	22.0	3.3	3.5	1.3	4.5
95-99	1,120	414,865	18.8	32.1	4.3	4.0	5.7	4.4
91-99	1,948	608,073	18.9	28.9	3.9	3.8	3.8	4.5
			Panel C: Low grade securities					
85-90	477	118,878	0.2	0.1	31.1	26.9		
91-94	820	191,901	29.1	38.0	18.2	15.6	24.3	21.4
95-99	1,175	376,570	19.9	24.1	9.5	9.8	12.0	12.6
91-99	1,995	568,471	23.7	28.8	12.8	11.5	18.2	16.5

Table 2

Securities issued and spread by type of underwriter, size, maturity, rating, type of issuer and sample periods

The table classifies the number and the total values of the securities in our sample – issued between January 1, 1991, and December 31, 1999, and with maturity longer than December 31, 1999 distinguishing between those underwritten by investment houses and by commercial banks, with a breakdown by size, original maturity, Moody's credit ratings, type of issuer and sample periods. Competitive issues are those made by issuers which used both banks and investment houses as underwriters during our sample period (Kang and Liu, 2007). Data on security issues are from Thomson Financial Database, defaults are from Moody's Default Report. "Bank issues" underwritings are those by one of the Section 20 security subsidiaries of commercial banks (i.e. deposit taking institutions with access to deposit insurance) listed in Cornett et al. (2002).

	Panel A: Investment house issues				Panel B: Bank issues				Panel C: Default rate	
	Non-defaulted issues		Defaulted issues		Non-defaulted issues		Defaulted issues		Investment houses	Banks
	Number	Spread *(basis point)*	Number	Spread *(basis point)*	Number	Spread *(basis point)*	Number	Spread *(basis point)*	Number	Number
Total	2,845	110	256	333	742	105	100	292	8.3	11.9
Size										
Less than $75m	615	97	37	163	45	90	4	286	5.7	8.2
$75m to $300m	1,653	135	163	380	471	104	65	325	9.0	12.1
Greater than $300m	577	135	56	229	226	138	31	222	8.8	12.1
Maturity										
Less than 5 years	505	75	31	135	106	65	11	75	5.8	9.4
5 to 15 years	1,219	195	118	431	350	146	59	364	8.8	14.4
10 to 20 years	739	145	75	350	171	125	23	309	9.2	11.9
Over 20 years	382	96	32	154	115	98	7	118	7.7	5.7
Rating										
Non-rated	85	632			7	578				
Aaa	51	45			12	34				
Aa1- Aa3	827	66	16	165	185	58	5	47	1.9	2.6
A1- A3	555	95	45	151	151	90	9	106	7.5	5.6
Baa1- Baa3	474	132	33	155	167	118	26	113	6.5	13.5
Ba1- Ba3	238	253	41	270	88	239	18	292	14.7	17.0
B1- B3	599	446	113	469	130	445	42	462	15.9	24.4
Caa1- Caa2	16	604	8	645	2	695			33.3	
Type of issue										
Investment house client	1,849	195	197	390					9.6	
Bank client					290	173	47	400		13.9
Competitive issue	996	84	59	155	452	95	53	123	5.6	10.5
Sample period										
1991-94	1,122	105	128	389	338	115	60	360	10.2	15.1
1995-99	1,723	134	128	241	404	113	40	138	6.9	9.0

Table 3

Effect of type of underwriter on security defaults

Panels A and B presents the results of probit regressions of the probability of default of securities underwritten by investment houses and commercial banks, where the dependent variable is a dummy that takes the value of 1 if the security defaults (equation 1 in the text). Panels C and D present hazard ratios. They are the results of a survival-time data model by the method of proportional hazards regression developed by Cox (1972), where the dependent variable is the 'baseline' hazard based on the duration measured in months (equation 2 in the text). For defaulted bonds the duration is the period from the date of issuance to the date of default. For non-defaulted bonds the duration is the period from the date of issuance to the date of repayment. Time and industry dummies are included, although not reported. Data sources for each variable are described in the notes to tables 1 and 2. High-grade ratings (i.e. HI-grade) are those classified by Moody's as having rating included between Aaa and Baa1. Low-grade ratings (i.e. LO-grade) are those classified by Moody's as Baa2 and below. Bank underwritings are those by one of the Section 20 subsidiaries listed in Cornett et al. (2002). Heteroskedasticity robust standard errors clustered at the issuer level are reported in parentheses. The symbol *** indicates a significance level of 1 per cent or less; ** between 1 and 5 per cent; * between 5 and 10 per cent.

	Panel A:		Panel B:		Panel C:		Panel D:	
Variables	Probit		Probit		Cox		Cox	
Bank Underwriting	0.201	*			1.389	**		
	(0.111)				*(0.226)*			
Bank Underwriting * HI-grade			-0.022				0.876	
			(0.210)				*(0.335)*	
Bank Underwriting * LO-grade			0.279	**			1.545	**
			(0.130)				*(0.284)*	
Size (log value)	0.022		0.026		1.046		1.054	
	(0.054)		(0.054)		*(0.094)*		*(0.093)*	
Gross spread	0.002	***	0.002	***	1.004	***	1.004	***
	(0.001)		*(0.001)*		*(0.001)*		*(0.001)*	
No. of Observations	3,279		3,279		3,569		3,569	
R-Square	0.273		0.274					

Table 4

Effect of type of underwriter on security defaults – Matching model

The table presents the results of a matching logit regressions of the probability of default of securities underwritten by investment houses and commercial banks (equation 3 in the text), splitting the sample by year of issue. The routine used for estimations is PSMATCH2, a Stata module developed by Leuven and Sianesi. (2003). The dependent variable is a dummy taking the value of 1 if the security defaults and 0 otherwise. The sources of the variable are described in the notes to tables 1 and 2. High-grade (i.e. HI-Grade) ratings are those classified between Aaa and Baa1 by Moody's. Low-grade ratings (i.e. LO-grade) are those classified by Moody's as Baa2 and below. Bank underwritings included as "treated" observations are those by one of the Section 20 subsidiaries listed in Cornett et al. (2002). The regression also includes calendar year, industry and rating dummies (not reported). Standard errors (reported in parentheses) are computed by bootstrapping, with 200 repetitions. The symbol *** indicates a significance level of 1 per cent or less; ** between 1 and 5 per cent; * between 5 and 10 per cent.

	Panel A: All Securities		
Variables	Treated	Controls	Difference
Average treatment effect on the treated	0.135	0.107	0.028** *(0.014)*
No. of Obs. (common support)	3,024	741	
No. of Observations	3,049	747	

	Panel B: LO-Grade Securities		
Variables	Treated	Controls	Difference
Average treatment effect on the treated	0.225	0.189	0.036 (0.025)
No. of Obs. (common support)	1,311	382	
No. of Observations	1,370	388	

	Panel C: HI-Grade Securities		
Variables	Treated	Controls	Difference
Average treatment effect on the treated	0.040	0.023	0.018* *(0.013)*
No. of Obs. (common support)	1,643	349	
No. of Observations	1,657	359	

Table 5

Effect of type of underwriter on security defaults – Split by sample period

Panel A reports the results of estimating the probit specification (equation 1 in the text) separating the bonds issued between 1991-94, while Panel B reports the results of estimating the probit specification (equation 1 in the text) separating the bonds issued between 1995-99. The column Difference test is the value of a test for the hypothesis that the coefficients of the regression are equal for the two subsamples, that is distributed as a χ^2 with 1 degree of freedom. Data sources for each variable are described in the notes to tables 1 and 2. High-grade ratings (i.e. HI-grade) are those classified by Moody's as having rating included between Aaa and Baa1. Low-grade ratings (i.e. LO-grade) are those classified by Moody's as Baa2 and below. Bank underwritings are those by one of the Section 20 subsidiaries listed in Cornett et al. (2002). Heteroskedasticity robust standard errors clustered at the issuer level are reported in parentheses. The symbol *** indicates a significance level of 1 per cent or less; ** between 1 and 5 per cent; * between 5 and 10 per cent.

	Panel A:	Panel B:	Panel C:
Variables	1991-94	1995-99	Difference test (first vs. second period)
Bank Underwriting	0.264 *	0.106	0.56
	(0.152)	(0.163)	
Size (log value)	-0.031	0.093	1.08
	(0.098)	*(0.069)*	
Gross spread	0.003 ***	0.002 ***	0.66
	(0.001)	*(0.001)*	
No. of observations	1,238	1,783	
R-Square	0.276	0.352	

	Panel D:	Panel E:	Panel F:
Variables	1991-94	1995-99	Difference test (first vs. second period)
Bank Underwriting * HI-grade	-0.365	0.142	1.97
	(0.380)	*(0.259)*	
Bank Underwriting * LO-grade	0.403 **	0.087	1.60
	(0.172)	(0.198)	
Size (log value)	-0.028	0.092	0.99
	(0.099)	(0.069)	
Gross spread	0.003 ***	0.002 ***	0.68
	(0.001)	(0.001)	
No. of observations	1,238	1,783	
R-Square	0.282	0.352	

Table 6

Effect of type of underwriter on security defaults – Split by type of issuer

Panels A and B report the results of estimating the probit specification (equation 1) of the paper separating the bond issues between competitive (those whose issuers have issued at least one bond underwritten by both investment houses and banks during the sample period) and uncompetitive. The column Difference test is the value of a test for the hypothesis that the coefficients of the regression are equal for the two subsamples. Data sources for each variable are described in the notes to tables 1 and 2. High-grade ratings (i.e. HI-grade) are those classified by Moody's as having rating included between Aaa and Baa1. Low-grade ratings (i.e. LO-grade) are those classified by Moody's as Baa2 and below. Bank underwritings are those by one of the Section 20 subsidiaries listed in Cornett et al. (2002). Heteroskedasticity robust standard errors clustered at the issuer level are reported in parentheses. The symbol *** indicates a significance level of 1 per cent or less; ** between 1 and 5 per cent; * between 5 and 10 per cent.

	Panel A:	Panel B:	Panel C:
Variables	Competitive issuer	Uncompetitive issuer	Difference test (competitive vs. uncompetitive)
Bank Underwriting	0.509 ***	0.068	4.10 **
	(0.136)	*(0.170)*	
Size (log value)	-0.009	0.185 ***	2.93 *
	(0.090)	*(0.069)*	
Gross spread	0.001	0.003 ***	1.52
	(0.001)	*(0.001)*	
No. of observations	1,068	1,923	
R-Square	0.279	0.323	

	Panel D:	Panel E:	Panel F:
Variables	Competitive issuer	Uncompetitive issuer	Difference test (competitive vs. uncompetitive)
Bank Underwriting * HI-grade	0.151	-0.185	0.39
	(0.215)	*(0.499)*	
Bank Underwriting * LO-grade	0.733 ***	0.110	5.81 **
	(0.183)	*(0.183)*	
Size (log value)	-0.008	0.191 ***	3.02 *
	(0.091)	*(0.070)*	
Gross spread	0.001	0.003 ***	1.77
	(0.001)	*(0.001)*	
No. of observations	1,068	1,923	
R-Square	0.286	0.335	

Table 7

Effect of Type of Underwriter on Security Defaults – Foreign banks and maturity

Panel A disentangles deals underwritten by foreign banks. Panel B describes the effect of maturity at issuance on defaults, including interaction terms for short (below 5 years), and long-term securities. For all panels the dependent variable is a dummy that takes the value of 1 if the security defaults (equation 1 in the text). Time, industry, maturity and rating dummies are included, although not reported. Variable sources are described in the notes to tables 1 and 2. Bank underwritings are those by one of the Section 20 subsidiaries listed in Cornett et al. (2002). Heteroskedasticity robust standard errors clustered at the issuer level are reported in parentheses. The symbol *** indicates a significance level of 1 per cent or less; ** between 1 and 5 per cent; * between 5 and 10 per cent.

	Panel A:		Panel B:	
Variables	Foreign		Maturity	
Bank Underwriting	0.188	*		
	(0.113)			
Bank underwriting * Foreign banks	0.995	***		
	(0.202)			
Bank Underwriting * short term			0.317	***
			(0.120)	
Bank Underwriting * long term			0.003	
			(0.166)	
Size (log value)	0.052		0.021	
	(0.051)		*(0.055)*	
Gross spread	0.002	***	0.002	***
	(0.001)		*(0.001)*	
No. of Observations	3,279		3,279	
R-Square	0.286		0.274	

Table 8

Effect of Type of Underwriter and deal characteristics on Spreads

Panels A to D present the results of the regressions of the regressions on the gross spreads of the security at issuance with respect to the benchmark with the probabilities of being underwritten by a bank (i.e. bank underwriting, dummy equal 1) or by investment houses (dummy equal 0). Maturity dummies for short, medium and long term maturity refer to maturities at issuance of below 5, between 5 and 10 and above 10 years respectively. High-grade ratings are those classified by Moody's between Aaa and Baa1. Low-grade ratings are those classified by Moody's as Baa2 and below. The dummies defaulted refers to all securities than eventually defaulted. Time, industry and rating dummies are included where applicable, although not reported. Variable sources are described in the notes to tables 1 and 2. Bank underwritings are those by one of the Section 20 subsidiaries listed in Cornett et al. (2002). Heteroskedasticity robust standard errors clustered at the issuer level are reported in parentheses. The symbol *** indicates a significance level of 1 per cent or less; ** between 1 and 5 per cent; * between 5 and 10 per cent.

Variables	Panel A	Panel B	Panel C	Panel D
Bank Underwriting	-1.577		-1.565	-2.364
	(3.242)		*(3.236)*	*(3.098)*
Bank Underwriting * HI-grade		-2.676		
		(2.490)		
Bank Underwriting * LO-grade		0.680		
		(5.668)		
Bank underwriting * Foreign banks			6.951	
			(11.180)	
Defaulted				28.899 ***
				(8.862)
Defaulted * Bank underwriting				-0.375
				(12.483)
Size (log value)	-0.538	-0.537	-0.400	-0.425
	(1.334)	*(1.335)*	*(1.357)*	*(1.358)*
No. of Observations	3,587	3,587	3,587	3,587
R-Square	0.841	0.841	0.841	0.841

Determinants of Collateral [*♥]

Gabriel Jiménez
(Banco de España)

Vicente Salas
(Universidad de Zaragoza and Banco de España)

Jesús Saurina
(Banco de España)

February 2004

Abstract

This paper draws upon the theory of contracting under asymmetric information to postulate and test several hypotheses concerning the relationship between the use and amount of collateral in financial loans to firms, the risk profile of the borrower, business cycle and monetary conditions of the economy, strength of the lending relation between borrower and lender, competition in the credit market and expertise and preferences of the lender. The research takes advantage of a very large panel of data coming from the Credit Register that contains the whole population of loans granted every year from 1984 to 2002 by Spanish banks to firms (approximately two million loans). Important novelties of the paper are that the quality of the borrower is measured in terms of *ex ante* and *ex post* credit risk, that the association between credit risk of the borrower and the use of collateral is evaluated in different segments of the credit market (short-term, and long-term loans, and new and old borrowers), which are likely to present differences in the relative information advantage of the borrower over the lender, and that we also control for borrowers' idiosyncratic effects. The evidence confirms that the use of collateral is determined in a predictable, different way in each market segment.

JEL: G21

Key words: collateral, asymmetric information, relationship banking, competition, business cycle

---

[*] This paper is the sole responsibility of its authors and the views presented here do not necessarily reflect those of the Bank of Spain. We thank everyone for the many comments and suggestions received at a CEMFI seminar, in particular, those made by M. Arellano, S. Bentolila, O. Bover and R. Repullo. We also thank A. Almazán for his comments. Any errors that remain are, however, entirely the authors' own.
[♥] Address for correspondence: Jesús Saurina; C/Alcalá 50, 28014 Madrid, Spain. Tlf: + 34 91 338 5080; e-mail: jsaurina@bde.es.

1. Introduction

Collateral is a common term of loan contracts, together with the interest rate, maturity, size and possible covenants. Credit markets research explains the use of collateral as a consequence of the problems of adverse selection (Besanko and Thakor, 1987a,b; Chang and Kanakas, 1985, and Bester, 1985), and/or the problems of moral hazard (Boot, Thakor and Udell, 1991), which are present in the transaction between borrowers and lenders. The presence or not of collateral can also be the result of more or less relationship banking (Sharpe, 1990, and Boot and Thakor, 1994), the result of the level of competition in loan markets (Besanko and Thakor, 1987a) or an alternative to a thorough screening of borrowers (Manove and Padilla, 1999 and 2001). Theories often make contradictory predictions on the relationship between the use of collateral and observed characteristics of the borrower, of the lender and of the relationship between the two of them. For example, the adverse-selection problem predicts a positive association between the use of collateral and quality of the borrower, while if the lender is informed about the quality of the borrower, the predicted association has a negative sign. Relationship-lending may evolve into a hold-up situation where the lender is able to impose stricter credit conditions on the borrower, including the condition to pledge collateral to get a loan, or in a situation where reputation concerns attenuate moral-hazard problems and reduce the demand of collateral.

Empirical research on the determinants of the use and amount of collateral in loan contracts will help to sort out the relative importance of the alternative-information conditions that affect transactions in credit markets, will clarify the net balance of the consequences of relationship lending and will provide evidence on to what extent lenders differ in their more or less thorough analysis of the loan operations. Furthermore, although the theories are so far silent about them, empirical research can provide further evidence on how macroeconomic conditions such as the business cycle and the real rate of interest affect the likelihood to use collateral in loan contracts. Over all, empirical research on the determinants of the use of collateral will provide a better understanding of how credit markets actually work, in the same way that we can learn about those markets from the determinants of the interest rates, maturity and covenants of loans.

The empirical evidence on why some loans have collateral and others do not is scant and gives contradictory results. Moreover, the empirical research has focused only on a limited set of issues such as collateral and credit risk (Berger and Udell, 1990, and Jiménez and Saurina, 2003) and collateral and relationship lending (Berger and Udell, 1995; Harhoff and Körting, 1998, and Degryse and van Cayseele, 2000). This paper contributes to the existing literature on the empirical determinants of collateral in loan contracts in several ways, all of which are related to the great advantage of being able to work with the whole population of financial loans annually made to Spanish business firms during the period 1984 – 2002. First, the paper tests, in a novel and more informative way, the relationship between the use of collateral and the quality of the borrower. Second, the paper evaluates the effect of relationship lending in loan contracts controlling for borrowers' specific effects. Third, the paper provides, for the first time, evidence on the likelihood of collateral depending on competitive conditions of the loans market, the business cycle, the real interest rate of the economy and characteristics of the lenders that reflect their relative propensity to substitute collateral for thorough credit screening. Finally, the paper also provides, for the first time, evidence on the determinants of the amount of collateral once the decision to use collateral is made.

One important research question in the literature on collateral is whether loans with collateral are riskier or not than those without collateral. The theory predicts that riskier borrowers are more likely to pledge collateral when the quality of the borrower is observable for the lender and there is moral hazard in the transaction (Berger and Udell, 1990, and Boot, Thakor and Udell, 1991). On the other hand, if the quality of the borrower is not observable (i.e., hidden information and adverse selection) then lenders will use collateral to self-select the borrowers according to their perceived risk and, in the equilibrium solution, high-quality borrowers will choose loans with collateral and lower interest rates (Bester, 1985; Besanko and Thakor, 1987a, and Berger and Udell, 1990). To find out which of the effects prevails, previous work has regressed the risk premium contained in the interest rate of the loan on the presence of collateral in the same loan (Berger and Udell, 1990, and Degryse and van Cayseele, 2000)[1].

[1] The empirical evidence is contradictory. For example, Berger and Udell (1990), in a sample of loans from the US credit market, find that the use of collateral is positively associated with higher risk premium

The interest rate and the use of collateral are jointly decided so therefore, the two are endogenous variables. This paper avoids the possible bias of the results due to this fact by using direct measures of the observed *ex ante* and *ex post* quality of the borrower, and testing the relation between risk and use of collateral in several samples of loans where differences in the relative importance of sorting by "private information" and sorting by "observed information" are expected. The testing of these information-related hypotheses using *ex post* quality of the borrower is, to our understanding, new in the literature on collateral[2]. For each loan in the sample, information is available on whether the borrower had a loan in default or not at the time the next loan is granted, and whether the borrower defaults on a loan in the year after the loan was granted, not having defaulted in the year before. Moreover, the total sample of loans in each year is divided into those loans that are made to "new" borrowers, in the sense that it is the first time they are given a loan by any bank, and the loans made to "old" borrowers. We expect that lenders know relatively less about the quality of the borrower in the cases where borrowers do not have a loan in default at the time they receive a loan, and in the cases where borrowers are new. Therefore the hypothesis to be tested is whether in these cases there will be relatively more borrowers sorted according to their private information than in the other samples and, consequently, one should expect a lower association between the use of collateral and quality of the borrower among borrowers that default on a loan in t+1, but not in t, and among new borrowers than among borrowers that are in default at the time of the loan and among old borrowers, respectively. The evidence confirms this prediction.

The methodology described above was possible because we work with the whole population of financial loans made to Spanish firms during the period 1984-2002, as is contained in the Credit Register database of the Bank of Spain. For each loan the Credit Register informs as to whether the loan has collateral or not and, in the affirmative case, if the collateral is partial, covers at least 50% of the loan, or total, i.e., it covers 100% of the loan. The pool of new loans made every year to business firms gives a time series

in the interest rate of the loan operation, while Degryse and van Cayseele (2000), from a sample of credit loans to small business by a Belgian bank, find a negative association between these two variables.

[2] Jiménez and Saurina (2003) test the relationship between the use of collateral and the risk of a loan's default using a direct measure of *ex post* credit risk, and therefore avoiding potential problems due to the fact that risk premium is jointly determined with the decision to use collateral. Harhoff and Körting (1998) also use the observed risk of the borrower at the time the loan is granted (whether the borrower is under financial distress or not) as a measure of quality, but they do not have information on *ex post* risk.

and cross-section database, which makes it possible to ask for the effects on loan conditions of exogenous variables common to all loans in a given year, such as the situation of the business cycle and the real interest rate.

This study also presents evidence on whether relationship-lending evolves into a hold-up situation of the borrower, in which case more relationship lending would imply more use of collateral (Sharpe, 1990) or, to the contrary, if it helps to overcome information asymmetries then stronger relations go together with less use of collateral (Boot and Thakor, 1994). Although previous evidence already exists on this topic, this evidence is inconclusive[3]. Our contribution to this issue comes from the panel of the population of loans used in the estimation, and from the fact that we have included more control variables, such as macroeconomic conditions and borrowers' specific effects in the empirical model. This is important because the variables of relationship banking are likely to be correlated with other unobserved characteristics of the borrowers; the fixed effects estimation of the model controls for such unobserved characteristics and avoids potential estimation bias of the relationship-lending effects on credit decisions.

The theoretical literature also views the decision to use collateral in loans as a substitute for a thorough evaluation of the operation by the lender (Manove and Padilla, 1999 and 2001). The hypothesis in this paper is that the substitution is more likely among banks that have less experience, resources and capabilities in the market of loans to business firms. One important issue related to this topic is whether savings banks, which are non-profit commercial banks that have substantially increased their share in the loan market during the period under study (up to fifty per cent), have a different likelihood than commercial banks for the use of collateral in business loans. It can be argued that savings banks will be more risk averse and have less expertise in the loan market than commercial banks, and this should translate into different lending policies.

Previous research has analysed differences in the determinants of the use of collateral by taking into account whether the collateral is internal or external to the loan operation

---

[3] In the case of a Belgian bank, Degryse and van Cayseele (2000) find that more relationship-lending implies more likelihood of collateral, while Berger and Udell (1995), for the US, and Harhoff and Körting (1998), for Germany, find that relationship-lending decreases the use of collateral.

(for example, Berger and Udell (1995) studied the differences between mortgages and accounts receivables or inventories). However, no previous work exists on the determinants of the *amount* of collateral used in a particular loan, even though there are theoretical results that provide predictions about this variable. We test the results of Boot, Thakor and Udell (1991), who show that under hidden information and/or hidden-action problems, the amount of collateral in a loan operation will increase with the real interest rate and will decrease with the size of the loan and with a reduction in the dissipative cost of the collateral.

Finally, the paper models the empirical determinants of the use of collateral for short-term loans (from one to three years of maturity) and for long-term loans (more than three years of maturity) separately. Previous research in this topic (Berger and Udell, 1995) has pointed out the relevance of the type of loan operation to explain the differences observed in the determinants of the term of the operation (interest rate and use of collateral). In this paper the analysis is limited to financial loans to business firms, but within them we observe that less than 8 per cent of the short-term loans have collateral, while this number is greater than 50 per cent for the long-term ones. The maturity of the loan severely conditions the probability to pledge collateral and we expect that the model that explains the decision to use collateral or not will be different in the two cases[4].

Although the data contained in the Credit Register is very rich, it is incomplete in several ways since nothing is known about the interest rate or the purpose of the loan, nor about characteristics of the borrower such as size (assets, sales or number of employees), profitability and financial leverage. As indicated above, the interest rate of the loan is another endogenous variable of the loan decision and will have to be explained in a separate model, different from the model of collateral. The rest of the variables are important as control variables and if they are not included in the model some of the results can be biased. To overcome this limitation, the paper introduces fixed effects for geographic and economic sectors of the lenders, as well as geographic and economic sector measures of credit risk, and includes the total amount of debt of the borrower at the end of the previous year among the explanatory variables, as a measure

[4] Maturity and size of the loan are considered additional explanatory variables of the use of collateral (Degrysse and van Cayseele, 2000).

of size. Moreover, when the robustness of the results is analysed, evidence is presented on the empirical determinants of collateral in a sub-sample of loans when controlling for borrowers' fixed effects. Except for the relationship-banking variables, whose effects on the use of collateral have to be evaluated again with the new evidence, the results obtained from the population of loans are quite robust.

Section 2 presents the literature review on the determinants of the use of collateral in loan contracts with a summary of the main empirical findings of previous research. In Section 3 we describe the database, the variables and the methodology followed to test the main hypotheses. The results of the empirical analysis are presented in Section 4, while Section 5 contains the discussion of the main empirical findings. The conclusions summarise the paper and highlight its contributions.

2. Literature review on the use of collateral in loan contracts and main hypotheses of the paper

The use of collateral in loan contracts has been explained in terms of the risk for the lender, in terms of the relationship over time between the lender and the borrower, in terms of the competitive conditions in the loan market and in terms of the relative efficiency of the lender when it comes time to evaluate the loan operation.

Collateral and risk

The first theoretical papers on the decision to pledge collateral consider a situation of asymmetric information at the time of the loan contract between the borrower and the lender, in the sense that the lender does not know the quality of the borrower (hidden information). The borrower's actions after receiving the loan are, however, public knowledge. In order to solve the adverse selection problem that emerges in these situations, Chan and Kanatas (1985), Bester (1985) and Besanko and Thakor (1987a) show that banks will offer to the potential borrower two different contracts to choose from: one with collateral and low interest rates, and the other one with no collateral and a higher interest rate. Those borrowers who consider themselves of high quality (low risk of default) will choose the contract with collateral, in order to take advantage of the lower interest rate. Borrowers of low quality will do otherwise. Borrowers sort themselves according to their private information. Under symmetric information, Besanko and Thakor (1987a) show that loan contracts would all be without collateral, but interest rates would increase with the estimated risk of the loan.

Conclusions change when the information asymmetry is in the form of hidden action, that is, the lender does not observe the action taken by the borrower after the loan is transformed into an investment project. Boot, Thakor and Udell (1991) show that with no hidden information, but with hidden action, the equilibrium loan decision implies that good-quality borrowers will get loans without collateral and bad-quality ones will get loans with collateral. The amount of collateral in each loan will be an increasing function of the real risk-less interest rate (in order to limit the increase in the cost of the loan and to cope with the moral hazard associated with higher interest rates), and a decreasing function of the dissipative cost of the collateral (since now the collateral is more worth to the bank and less collateral is needed to break even) and of the size of the loan (since larger projects increase the payoff to the borrower in good states). Finally, under both, hidden action and hidden information, in general, most of the loans will be with collateral, but the amount of collateral in each of them will increase with the real risk-less interest rate (Boot, Thakor and Udell, 1991).

The first empirical formal tests of these predictions correlate the risk premium charged in the loan operation with the presence of collateral in the same operation (Berger and Udell, 1990). On the one hand, if collateral helps to solve problems of adverse selection then the risk premium of the loan should be negatively correlated with the likelihood of collateral since the theory predicts that collateral and lower interest rates go together in the package offered by the lender to sort the borrowers. On the other hand, if lenders know the quality of the borrowers, then a positive association should be observed between risk premium and the presence of collateral in the loan, since the theory predicts that in these conditions lenders will ask low-quality borrowers for collateral and higher interest rates, and lower interest and no collateral from the high-quality, low-risk ones. Berger and Udell (1990) find a positive association between use of collateral and risk premium which, would be consistent with the assumption that lenders sort the borrowers from the information they have about their quality (observed-risk hypothesis)[5], but Degryse and Van Cayseele (2000), in a sample of loans from a single bank, find the opposite result, that is, risk premium is lower in collateralised loans.

Jimenez and Saurina (2003) test the relationship between the use of collateral and credit risk, explaining the probability of default for a particular loan as a function, among other variables, of whether the loan has collateral or not. They find that the probability of default (a measure of *ex post* credit risk, instead of the *ex ante* risk premium measure) is higher in loans with collateral, which would be consistent with the sorting by the observed-risk hypothesis at the time the loan decision was made.

In this paper a distinction is made between risk related with the quality of the borrower (idiosyncratic risk) and risk from the business cycle and economic conditions of the local market. The quality of the borrower is measured from the information available in the Spanish Credit Register data-base in terms of whether the borrower has a previous loan in default or not at the time the loan decision is made, and if the borrower who gets the loan defaults the following year or not. If the borrower has a loan in default at the time of receiving a new one, then she can be considered a low-quality borrower. Since the information is available to the lender, even if the borrower is new to the bank (but was already in the system), a positive association between having a defaulted loan and the use of collateral may be expected from those sorted by the observed information paradigm[6].

The credit quality of the borrower in the following period after the loan was made (measured by the indication of whether the borrower which had not defaulted before, defaults or not the next year) should be negatively correlated with the use of collateral at the time of the loan decision if sorting is made under the private information assumption, and positively correlated under the observed quality of the borrower assumption. The fact that we can identify those borrowers who default in the time-period following the loan decision’s being evaluated allows for a clearer test of the

[5] The authors consider their results consistent with the conventional wisdom in the credit markets about collateral and risk, and with previous empirical work that in a less formal way found that loans with collateral were classified as riskier by the bank analysts (Orgler,1970, and Hester, 1979).

sorting by observed risk or the sorting by the hidden-information hypothesis than in other papers, where the quality of the borrower before and after the loan operation is not available.

For those borrowers who get a loan for the first time, the information available to the lenders, in order to evaluate their respective quality, will be limited. The information asymmetry between a borrower and a lender will be more pronounced than in the case where the borrower has been in the system for a long time. We would expect that for new borrowers the situation of hidden information about the quality of the borrower will be more frequent than for the rest of borrowers and, consequently, the information asymmetry is likely to be solved in the way predicted by the sorting by private information models. If this were the case, the positive association between collateral and risk of the lender would be milder among the new borrowers than among the old ones, or even negative if the sorted by private information model prevails.

The information asymmetries between lenders and borrowers not only affect the decision to use collateral but also the amount of collateral (Boot, Thakor and Udell, 1991). The paper provides empirical evidence on whether the amount of collateral increases with real interest rate of the economy and decreases with the size of the loan, as the theory predicts.

No theory is available on how macroeconomic conditions affect the use of collateral in loans. Presumably, lenders will anticipate that, in periods of expansion, the probability of firms entering into some financial distress and, therefore, of having problems to return the loan is lower than in periods of contraction or even recession. Each lender may fear that her loan will be unsecured if others anticipate riskier economic conditions because of lower growth and decide to ask for collateral. The impact of the business cycle may induce lenders to increase the demand of collateral in periods of lower economic growth and reduce it in periods of expansion. The monetary conditions of the economy (high or low real interest rates) may also affect the observed use of collateral in loan decisions: under tighter monetary conditions (higher interest rates), lenders

[6] Harhoff and Köring (1998) find that the probability of collateral is significantly higher among the lines of credit of firms that are under financial distress at the time the loan decision was made than among those that were not.

give priority to their best borrowers[7]. Under this assumption an inverse relation between real interest rate and use of collateral will be expected. Since many borrowers limit their activities to a particular geographic market, besides the macroeconomic conditions common to all of them, lenders may take into account the state of the economy in the local markets and/or in the economic sector of the borrower.

Relationship lending

Banks and firms often maintain many relations repeated over time. The scope and strength of the relations generate private and customer-specific information and create conditions for what is called "relationship banking" (Boot, 2000), quite different from "transaction-based" banking (Berger and Udell, 2002), where lenders and borrowers transact under an arms-length contract. Relationship banking allows lenders to learn about the hidden attributes and hidden actions of the borrowers and, as a result, decisions such as the availability of credit, the interest rate on the loans, and to pledge collateral or not, may be done taking into account the expected length of the relationship. Diamond (1991) and Petersen and Rajan (1994) present models in which interest rates of the loans decrease over time as lenders learn more about the attributes and decisions of the borrowers. Repeated interactions also favour implicit contracts sustained by reputation considerations with consequences for the loan conditions. In this respect, Boot and Thakor (1994) present a model where collateral requirements are related to the length of the relationship. Borrowers pledge collateral early in the relationship but do not pledge collateral later, after they have demonstrated some project success[8]. Therefore, a negative relation should be expected between the use of collateral and the length of the bank-firm relationship.

Another possible consequence of relationship banking comes when it ends up becoming a "hold-up" situation, i.e., the information advantage of the banks may evolve into a monopoly position. If this were the case, the terms of the loan contract could worsen over time, including the need to pledge collateral in order to have access to loans

[7] Notice that this prediction is on the decision to pledge collateral or not, and not on the amount of collateral being used.
[8] Internal collateral will only be effective if its value can be monitored (Rajan and Winton, 1995). Relationship-banking may facilitate the monitoring of pledged collateral and, therefore, it will facilitate the use of internal collateral.

(Greenbaum et al., 1989; Sharpe, 1990, and Rajan, 1992). Under this assumption, a positive association between collateral and the length of the bank-firm relationship would be expected.

Berger and Udell (1995) for the US, and Harhoff and Körting (1998) for Germany provide evidence consistent with the models of relationship-banking and reputation that predict an inverse relation between intensity of the bank-firm relationship and use of collateral. But Degryse and Van Cayseele (2000) find a positive association between a measure of relationship-banking and the probability of collateral in a loan operation. The empirical results in previous work are, therefore, contradictory, although the first two papers are limited to lines of credit loans and the third covers all kinds of loans but limited to SMEs exposures of one bank.

In this paper the strength of the bank-firm relationship is measured by the duration of the relation between the borrower and the bank that gives the loan. We expect this variable to have a negative coefficient in the equation that estimates the probability that the loan is secured with collateral if reputation and/or learning considerations affect the loans market, and a positive one if longer relationships evolve into hold-up situations. Other variables that also capture part of the causes or consequences of relationship-lending, such as the number of loans that the lender has with the borrower at the beginning of the period, and the number of banks with which the borrower has loans, also at the beginning of the period, are also included in the model. As the number of loan operations with the bank increases, the bank will accumulate more knowledge about the borrower and the reputation effects may be stronger so that the variable is likely to be negatively correlated with the use of collateral. On the other hand, a bank may force the borrower to fraction the total debt into a number of small loans, because it has some doubts about her quality. The effect of the variable number of loans with the lender in the use of collateral in a new loan is uncertain.

As the number of banks the firm works with increases, the bargaining power of the borrower, with respect to each of the banks, will increase and the lender is more likely to obtain loans without collateral. In other words, more lenders reduce the possibility of hold-up. At the same time, if the debt is shared among many lenders, controlling for the total amount borrowed, each lender will have more difficult access to the assets of the firm, and to ask for collateral is a way to protect the loan. Under this assumption, a positive relation between number of lenders and use of collateral would be expected (Harhoff and Körting, 1998). The final net result is once more an empirical question.

Part of the contradictory empirical results obtained in models that explain the conditions of the loans are determined by the different type of loans that are contracted. Berger and Udell (1995) explain their contradictory results, along with those of Petersen and Rajan (1994), by the argument that the line of credit, the type of credit considered in their analysis, is a type of loan more suitable for the benefits of relationship-lending. In this paper, the model on the determinants of collateral in loan operations is estimated separately for short-term loans (those with maturity of up to three years) and for long-term loans (more than three years of maturity). This is made to assure that the loan operations for which we study the determinants of collateral are more homogeneous. As indicated above, only 8 per cent of the short-term loans have collateral, while the percentage of loans with collateral among the long-term ones is over 50. More likely, as the maturity of the loan increases the information advantage of the borrower over the lender is likely to increase in events that may occur later in the future.

This way, in long-term loans the information advantage of the borrower over the lender, about the risk of the operation, is likely to be more pronounced than in short-term loans, where the nature of the relationship is also shorter. In other words, banks can use past information to anticipate what may happen in the next one or two years, but it may be more difficult for them to look forward beyond this time-period. If this is true, in long-term loans there are more opportunities for borrowers to signal their (future) quality to the lenders. In terms of the empirical model, the relation between collateral and risk should be milder in long-term loans than in short-term ones, since in the former the information asymmetry between the two is more pronounced than in the latter.

Competition in the loans market

More or less competition in the loans market also determines the terms and conditions of the individual loan contracts. Besanko and Thakor (1987a) show that in the absence of hidden-action problems, a monopolistic lender will never ask for collateral in loan operations. On the other hand, when the loan market is competitive, then hidden-information problems (but not moral-hazard) will give equilibrium solutions in which good borrowers will choose loans with collateral and lower interest, while bad borrowers will choose loans with higher interest but no collateral. Therefore, from these results the likelihood of collateral in a loan operation is higher in more competitive markets.

Moreover, competition will also have an indirect effect on loan conditions since it is often recognised that the development of relationship-banking will depend upon the intensity of such competition (Chan et al., 1986; Diamond, 1991, and Petersen and Rajan, 1995). Under more competitive conditions bank-firm relationships will have more difficulty in spreading the benefits of a single operation over time, the loans will necessarily be of shorter maturity and the incentives to invest in information will diminish. Relations will be weaker and shorter, and therefore reputation considerations will be less relevant at the time to decide to ask for collateral or not. From these arguments the probability of collateral should be higher in more competitive markets since to pledge collateral may be the way to secure the loan when there are no incentives to invest in information to evaluate the quality of the lender, and the reputation of the borrower has lower economic value.

No previous empirical work on the determinants of the use of collateral in loan operations has tested for the relationship between competition in the loan market and the presence of collateral. In this paper the province where the borrower contracts the loan as a separate regional market and measures the level of concentration in the province loan market is considered. Since less competition is expected in more-concentrated markets than in less-concentrated ones, a negative association is expected between market concentration and use of collateral.

Experience and preferences of the lender

To ask for collateral as a condition to get a loan may be considered by the lender as an alternative to screen and evaluate the actual risk of the borrower, and of the particular loan operation (Manove and Padilla, 1999 and 2001). Banks with a lower level of expertise (in particular, new entrants in product or regional markets) and with scarce resources to evaluate the economic risk of the loan will have more incentives to use collateral as a substitute for such evaluation. Some papers have looked at differences in lending practices among different financial intermediaries, taking into account their specialisation (Carey et al., 1998) and their ownership structure (Saunders et al., 1990, and Gorton and Rosen, 1995), but the relationship between type of lender and the collateral decision has not been studied in detail.

Banks in the sample have different specialisation in financial loans for business firms, different size and some of them are commercial banks while others are savings banks. Banks more specialised in financial loans and banks of larger size should have a comparative advantage in terms of evaluating the risk of borrowers, and one should expect a lower recourse to the collateral solution than in the rest of the banks. Savings banks (non-profit commercial banks) have traditionally been more concentrated in loans to households and only more recently have expanded into loans for business firms. Moreover, savings banks are more conservative institutions as they are, in practice, under the control of their managers and workers interested in keeping their jobs and avoiding possible financial distress. If this is true, the probability of collateral in a loan should be higher among savings banks than among commercial banks.

3. Database, variables and methodology

Database

The database used in this paper is the Credit Register of the Bank of Spain (Central de Información de Riesgos, CIR). This database records monthly information on all loans granted by all credit institutions (commercial banks, savings banks, co-operatives and credit finance establishments) in Spain for a value of over 6,000 euros (one million pesetas before the advent of the euro).

The CIR includes information on the characteristics of each loan (instrument, currency, maturity, amount of collateral, default situation and amount drawn or available) and of each borrower, both legal (i.e., firms) and "natural" persons (i.e., individuals): the province in which they operate their business and the economic sector in the case of legal persons and individual businessmen. In this work all the variables, except those related to macroeconomic conditions, have been obtained from the CIR.

We focus on all financial loans above 6,000 euros granted by any Spanish commercial and savings bank to legal persons during the time-period between December 1984 and December 2002. Collateralised loans to individuals are mainly mortgages, where the collateral is internal to the operation (it is very rare to get a loan to buy a house where the house is not pledged as collateral). Furthermore, the loan to buy the house is transaction-driven and we want to focus on operations that can evolve into relationship-banking. The reason to concentrate only on financial loans, and exclude other debt operations such as commercial loans, leasing, factoring operations as well as off-balance sheet commitments, responds to homogeneity reasons and to the belief that they are more suitable for relationship-banking. Nevertheless, financial loans are the bulk of loans to firms[9]. Finally, commercial and savings banks concentrate 95 per cent of the financial loans, and the comparison between the two forms of ownership is easier to interpret.

In this large database there is not a code for each loan. Thus, it becomes very difficult to follow a loan over time. Since we wish to know the precise year when the loan is granted, the sample finally selected has been limited to those loans for which it was possible to estimate the year when it was made. For example, loans with a maturity shorter than one year have been excluded because among them it is impossible to know if they are new or the result of a past loan that is financed again. As a result of this process the database contains, for each December between 1985 and 2002, all loans granted during the year with a maturity of one year or more.

Loans in the database are split into two groups: short-term (one to three years of maturity) and long-term (more than three years). As indicated above, this separation responds to the purpose of having more homogenous operations within each of the two samples, taking into account the observed differences in the frequency with which collateral is used in each of the samples (8 and 50 per cent, respectively). Unfortunately, we do not have information on the nature of the investment project financed by the loan, or on whether the loan is a line of credit or not.

[9] Financial loans are 60% of the total number of loans in the CIR.

A second classification of the loans has to do with the experience of the borrower. One group of loans contains those loans from borrowers about whom, at the time the loan is granted, there is already past information in the CIR database. The other group of loans, which we call new ones, are from borrowers who appear in the CIR database for the first time.

Variables and expected effects on collateral

Table 1 describes the variables used in this study. We have grouped them into six classes taking into account whether they are dependent or explanatory variables, and among the latter group depending upon whether they are control variables or belong to one or another theory about the determinants of the use of collateral: borrower's risk, economic conditions, relationship-lending, competition, and efficiency and preferences of the lender.

The dependent variable COLLATERAL is a 0/1 variable that takes the value of 1 if the loan has any kind of collateral, and 0 otherwise. The amount of collateral is measured by a variable that takes the value of 1 if the loan has collateral that covers more than 50 per cent, but less than 100 per cent of the risk, 2 if the loan is 100 per cent secured, and 0 otherwise.

The *risk of the borrower* is approximated by a list of variables that refer to the quality of the borrower. The variable DEFAULT indicates if the borrower had any loan defaulted in the previous year when the loan was made (t-1), and if the borrower did not have a defaulted loan in the previous year but does have a loan defaulted in the year after the loan was granted (t+1). The historic record of any borrower in terms of fulfilling the debt obligations or not with the lenders is made available to all lenders by the CIR. Therefore, this variable, at t-1, is common knowledge and it is a very good proxy of the variable quality of the borrower considered in the theoretical models. The empirical model would predict a positive association between the use of collateral and DEFAULT (t-1) from the sorting by the observed-risk paradigm.

Ex post risk of the borrower reflects the information conditions at the time the loan decision is made. If the lender could observe the quality of the borrower at the time of the decision then among those borrowers that default *ex post* the likelihood to have collateral will be higher than among those that do not default. In other words, the variable DEFAULT (t+l) should be positively associated with the use of collateral. If lenders can not observe the quality of the borrower at the time of the decision, the signalling-equilibrium model predicts that good-quality borrowers will choose loans with collateral. Therefore, a negative association between DEFAULT (t+1) and the use of collateral will indicate that borrowers were sorted according to the private-information paradigm.

More likely, some loans will be made under one type of sorting condition, and others with the alternative one. What we observe is the net effect and the positive or negative signs indicate which one dominates in the sample. The sorting by the private-information paradigm is expected be more frequent among new borrowers, and possibly among long-term loans, where the borrower can know more than the lender about the situation of the firm several years into the future. From this assumption the empirical prediction would be that the increase in the likelihood to pledge collateral among those borrowers that default in t+1, compared with the likelihood for those that do not default, is smaller among the new borrowers than among the older ones, and smaller in long-term loans than among short-term loans.

The total amount of debt of the borrower at the start of the year, DEBT (t-1), is a measure of risk for the lender since higher debt, *ceteris paribus*, implies more probability of financial distress. But more DEBT may also imply that the borrower is a larger firm for whom less probability of financial distress is expected. Therefore, the empirical association between DEBT and the use of collateral will have to be interpreted taking into account that the variable may be correlated with the size of the borrowers.

Older firms are often viewed as less risky since they have been able to survive under the threat of competition for a longer period. Age also signals experience and the opportunity by the lender to learn more about the hidden capabilities and assets of the borrower. “AGE” AS BORROWER (t-1) is equal to the number of years since the borrower had the first loan operation registered in the CIR (the year the firm was founded is not available). It is assumed that this “age” is appropriate as an inverse measure of risk of the borrower since her presence in the CIR determines the opportunities to learn about the borrower by the lenders. Notice also that “AGE” AS BORROWER is also a control variable to properly evaluate the effect of relationship-lending in the loan decision, since the number of years a borrower is associated with a bank is relatively long or short depending on the age of the borrower.

The macroeconomic environment at the time of the loan decision is captured by the variables GDP GROWTH, contemporaneous and lagged, and by the PHASE OF THE ECONOMIC CYCLE, to try to capture any asymmetry (absolute value of the difference between GDP GROWTH and the average of the variable for all of the years in the sample) and REAL INTERST RATE. The macroeconomic risk should be lower when the economy grows at a faster rate and when the economy is in an upturn, so a negative association is expected between the business-cycle variables and the use of collateral. As indicated in the theory section, high real interest rates, controlling for growth, may indicate tighter monetary conditions and more restrictions in the availability of credit; in these circumstances banks will give priority to their preferred customers, who are also the better ones, and this preference should translate into a negative association between interest rate and the use of collateral.

The risk profile of the province where the borrower is located, GEOGRAPHIC RISK, is measured by the proportion of bad loans over total loans in the province in the previous year. A similar variable, the proportion of bad loans over the total loans in the economic sector of the activity of the borrower, up to 11, is PRODUCT MARKET RISK, which captures the different economic conditions that may be present in the product market of the borrower.

The variables of *relationship-lending* considered in the paper are: the number of years since the borrower had the first loan with the current lender DURATION (t-1); number of loan operations alive with the borrower at the time the loan was made, SCOPE (t-1), and number of lenders with which the borrower has relations, NUMBER OF LENDERS (t-1). The variable DURATION captures the private information revealed through the intermediation process, and from the positive side of the relationship-banking literature a negative association is expected between this variable and the use of collateral. The sign of the relation would be the reverse if longer relationships create hold-up problems.

From the theoretical discussion, the effect of the variables SCOPE and NUMBER OF LENDERS in the likelihood of the use of collateral is also uncertain depending on whether relationship-lending evolves into more trust and less moral hazard or into a hold-up situation. Another important issue to account for is that relationship-lending variables are among the few variables that capture characteristics of the borrower. This means that if they are correlated with other unobserved variables, for example, the total assets of the borrower, then their estimated coefficients can be biased. The robustness test performed later will be of great help to properly evaluate the effects of the relationship-lending variables in the use of collateral.

It is assumed that *competition* among lenders is more intense in less-concentrated markets. Each of the 50 provinces of Spain is considered a separate geographic market and concentration is measured by the HERFINDHAL index, equal to the sum of squared shares of loans for all banks that are present in the province. A negative association is expected between market concentration and use of collateral.

The *efficiency, preferences and specialisation* of the lenders are approximated by the variable SIZE OF THE BANK: assets in relative terms to the total assets of the banking system; SPECIALISATION: proportion of loans to legal persons; and type of ownership, where the variable BANK takes the value of one if the lender is a commercial bank and zero if it is a saving bank (non-profit commercial bank). The theory predicts a negative association between all of these variables and the use of collateral.

Finally, the control variables are SIZE OF THE LOAN and two dummy variables, PROVINCE and ECONOMIC SECTOR, which take the value of 1 if the loan is made in one of the 50 provinces or in one of the 11 economic sectors identified by the CIR, respectively. The size of the loan may also be considered a proxy of risk since, controlling for the total debt before this new loan, larger loan size implies a higher increment in the leverage position of the borrower. So SIZE OF THE LOAN is positively associated with the use of collateral. But once the decision to use collateral is made, Boot et al. (1991) predict that the amount of collateral decreases with the size of the loan; this prediction will also be empirically tested.

The descriptive statistics of the data are presented in Table 2, for “old” borrowers, and Table 3, for the “new” borrowers. Notice first that, as indicated, the presence of collateral is much more frequent in long-term loans, over 50 per cent have collateral, than in short-term loans, only 8-9 per cent have collateral. The proportion of loans with collateral is slightly higher, 10 per cent, in new borrowers than in old borrowers.

The average size of the loan in the short-term group is half of the size as in the long-term one, as could be expected by taking into account that the short-term loans can be renewed along time. But notice also that the average of the variable DEBT, which proxies the size of the borrower, is also twice as large in the column of long-term loans than in the column of short-term ones, which implies that long-term loans are more frequent among large firms (borrowers) than among small firms. The ratio between the average size of the loan among long-term loans and among short-term loans is the same for old and new borrowers, although the average size of the loans among the former is three times the size of the latter. This means that new borrowers start to borrow with relatively small loans, probably because the firms in the group are smaller on average than the older ones or, perhaps, the bank is more cautious. Besides, the standard deviation shows that the size of the loan is more homogeneous in the short-term than in the long-term, and much more for new borrowers than for old ones.

Larger banks tend to make more long-term loans than smaller ones, as the average size of banks among the group of long-term loans is larger than the average size among the short-term ones. The average lender of the new borrowers is smaller than the average lender of the old borrowers, probably because in general small (new) firms tend to do more business with small banks while large firms do business with larger customers (Berger and Udell, 1995, and Peek and Rosengren, 1996). Commercial banks have a larger market share than savings banks in the short-term loans, among both old and new borrowers. Besides, the market share of the savings banks is higher among the new borrowers than among the old ones, probably because savings banks have increased the market share in the financial loans market over time getting a larger share of the new entrants, and among smaller firms in general, than commercial banks, while lending longer-term.

The variables of relationship-lending and the variable of AGE AS A BORROWER show very similar values in the two columns of short- and long-term loans. These variables present a big dispersion in the database, especially SCOPE and NUMBER OF LENDERS. The very high "maximum" values of these variables correspond to borrowers that do business in many provinces and have loan contracts with different banks in each province.

The proportion of borrowers that have a loan in default at the time they get a new loan, DEFAULTED (t-1), is between 3 and 4 per cent for old borrowers. Among those borrowers who are not in default when they get a short-term loan, 11 % have defaulted in the following year; this percentage is 9% for longer-term loans. Among new borrowers, 17% of those with a short-term loan will have defaulted in the following year after they got the loan, while the percentage of defaulted loans in t+1 among the new borrowers with a long-term loan is 9%. New borrowers with short-term loans have higher *ex post* credit risk than old borrowers with loans of the same maturity, although for longer ones this credit risk is the same for both groups. Since the proportion of loans with collateral is similar in the groups of old and new borrowers, the preliminary evidence does not support the hypothesis that loans with collateral will be more frequent among high-risk borrowers than among low-risk ones, at least for long-term loans (Section 5 analyses this fact in-depth.)[10].

Methodology

The hypothesis about the determinants of the use of collateral in financial loans will be tested from the econometric estimation of a Probit model, which explains the dependent variable COLLATERAL as a function of the explanatory variables listed in Table 1.

[10] It may be of interest to compare some of these descriptive statistics with those presented by Berger and Udell (1995), for the US, and by Harhoff and Körting (1998), for Germany, keeping in mind that they work with credit-line data. In the US sample 53% of the loans have collateral while in the German sample the proportion is 62.4%. In the US sample borrowers have been with the bank they get the line of credit from 78% of the time since they started the activity (11.4 years of Duration and 14.1 years of Age), while for the German sample the percentage is 70% (12 and 17.5 years). In our sample the "age" of the borrowers starts to count at the time when each one gets the first loan (upon appearing in the Credit Register for the first time), and the ratio of loyalty to the lender is 52% (2.95 and 5.66 years). Finally, almost 30% of the firms in the German sample had been in financial distress during the past 5 years, (3% of our firms had a loan in default the year before they got a new loan) while the average number of lenders they have relations with is 1.82 (3.94 in our sample). The German and US samples consider only small firms, while we have the population of Spanish firms as well as the population of loans to them.

The estimation will follow the "two-stage conditional maximum likelihood" (2SCML) procedure, developed by Rivers and Vuong (1988), to take into account that the SIZE variable of the loan can be endogenous, i.e., jointly determined with the decision to ask for collateral.

Our model is composed of a structural equation of the Probit model that is of primary interest, and another reduced-form equation for the endogenous explanatory variable. The method works this way: in the first step, the endogenous variable (log of the size of the loan) is written as a function of all of the exogenous explanatory variables. The model is estimated by ordinary least squares and its standardised residuals are considered as a new explanatory variable of the use of collateral. In the second stage the Probit model is estimated excluding the variables chosen as instruments of the SIZE variable and including the standardised error term obtained in the first step.

The variable chosen as instrument has been PRODUCT MARKET RISK, i.e., the proportion of bad loans over total loans in each of the eleven economic sectors lagged one period. This variable is highly significant statistically (with a negative coefficient) in the SIZE equation and, empirically, is not correlated with the COLLATERAL variable. On the one hand, the proportion of non-performing loans in an economic sector provides an indication of the opportunities for long-term investment as lenders are more willing to finance large projects in those sectors where firms have fewer difficulties to repay the loans. Finally, lenders may take into account the economic risk of the loan in determining the interest rate but not in the decision to use collateral or not.[11]

Under this methodology, the parameters of the structural equation are not those obtained in the second-stage Probit model and have to be recovered with a simple transformation. Denoting by $\acute{l}$ the estimated parameter for the variable "first-stage standardised error term", the original parameters for all the other explanatory variables in the Probit model will be equal to the estimated one multiplied by the term $1/\sqrt{1+\acute{l}^2}$.

[11] We have also tested with other instruments such as DEBT (t-1), and the main results remain the same.

This procedure also allows for the statistical testing of the hypothesis of endogenous variable for SIZE of the loan. To do so, it compares the log of the likelihood functions of the Probit model calculated with ($\ln Lu$) and without ($\ln Lr$) the standardised residuals obtained in the first stage. This is a widely used likelihood ratio test where the null hypothesis is that the variable is exogenous. The statistic:

$$LR = -2(\ln Lr - \ln Lu),$$

follows a Chi-square distribution with degrees of freedom equal to the number of endogenous variables in the Probit equation[12].

The estimation of the Probit model is repeated, following the procedure described above, for short-term and long-term loans, as well as for the two sub-samples of old and new borrowers.

Finally, the basic model on the determinants of the use of collateral is estimated for a sub-sample of loans to old borrowers in which it is possible to control for borrowers' specific effects. The new estimation is introduced into the analysis as a robustness test because the Logit model with fixed effects can not be applied to the whole sample, but it is an important methodological contribution since it is a reliable way to find out if the omission of variables, correlated with these effects in the general model, introduces biases or not into the estimations.

The other hypotheses have to do with the determinants of the amount of collateral. The dependent variable distinguishes between loans with an amount of collateral over 50% and loans with 100% collateral. The prediction is that the likelihood of using 100% collateral relative to the likelihood of using partial collateral increases with the size of the loan and with the real interest rate. To perform the test will require the estimation of a multinomial Logit model[13].

[12] Since, in our case, the number of endogenous variables is one, the LR test and the significance of the error term of the loan size are equivalent, so we will refer to both indistinctly.

[13] DEBT, AGE AS BORROWER, DURATION, SCOPE, NUMBER OF LENDERS and SIZE OF THE LOAN will enter into the Probit and into the multinomial Logit models in natural logs because we expect decreasing marginal effects on their effects in the probability of use of collateral in a particular loan. Also, most of the continuous variables are measured at the beginning of the period to avoid simultaneity problems. DEBT and SIZE OF THE LOAN are in constant prices of the year 2002. It is implicitly

4. Empirical Results

4.1. The use of Collateral

Tables 4 and 5 show the results obtained from the estimation of the Probit model for old and new borrowers, respectively. Each table presents the results for short-term and for long-term loans. Besides, to better evaluate the economic relevance of the results, Tables 4 and 5 are complemented with Table 6, which gives the marginal effects on the probability of using collateral in a financial loan to marginal changes in the explanatory variables.

- *Old borrowers*

The Probit models are jointly statistically significant in the two groups of loans. The statistical tests confirm that SIZE of the loan is an endogenous variable [14]. The variable DEBT is the only one not statistically significant in the two groups of loans (this variable explains the size of the loan but not the presence of collateral), while GEOGRAPHIC RISK and SCOPE are not significant in the group of long-term loans.
Overall, the pattern of results is quite similar in the two groups of loans and in line with the prior hypothesis. The use of collateral increases with the risk of the borrower, both *ex ante* (the coefficient of DEFAULTED (t-1) is positive) and *ex post* (positive coefficient of DEFAULTED (t+1)), while it decreases with the "AGE" of the borrower (negative coefficient).

The downturn of the economy increases the likelihood of the use of collateral (negative coefficient of GDP GROWTH). In fact, a more thorough analysis shows that output growth has an asymmetric effect on the use of collateral: we observed a different impact depending on whether output growth is above its average or whether it is below its

---

assumed that the multinomial Logit model is a good approximation of the Probit model, which is the proper one when one of the variables is endogenous.

[14] Not controlling for the endogeneity problem, the coefficient of the SIZE of the loan variable in the group of long-term loans would have been twice the value obtained with the correction (0.192+0.202, instead of 0.192), while in the group of short-term loans the correct parameter would have been underestimated by 40% (-0.154/0.384).

average (captured by the PHASE ECONOMIC CYCLE). In this way, the likelihood of collateral diminishes proportionally more in the former situation than it increases in periods when actual growth is less than the long-term average. For instance, in the long-term (short-term), when GDP growth is below its average, the global effect is -0.09 (-0.05) compared with the -0.06 (0.00) in the upturns.. In periods of tighter monetary policy or higher real interest rates, the use of collateral is less likely than in periods of loose monetary policy (coefficient of REAL INTEREST RATE is negative).

More relationship-lending decreases the likelihood of the use of collateral (negative coefficient of DURATION and positive of SCOPE), while to have loans with a higher NUMBER OF LENDERS seems to increase the bargaining power of the borrower, with the result of lower use of collateral (negative sign of the coefficient).

In more concentrated markets (less competition) the likelihood of the use of collateral is lower than in less-concentrated ones (negative sign of the coefficient of HERFINDHAL).

Finally, more experienced and specialised banks tend to decrease the demand of collateral in their loan operations (the variables SIZE OF THE LENDER, BANK and SPECIALISATION all have a coefficient with a negative sign).

From Table 6, to be a low-quality borrower at the time of getting a new loan increases the likelihood of having to pledge collateral, with respect to those borrowers who do not have any loan defaulted twice for a new loan with long maturity than for a short-term loan (0.148 compared to 0.078). The marginal effect of being a borrower who has defaulted a loan in t+1, but who did not have any loan defaulted in t-1, is similar in the two groups of loans and smaller than in the case of having defaulted in t-1.

The marginal effect of the business-cycle variables is overall higher in the long-term loans, i.e., a point reduction in the lagged GDP GROWTH increases the likelihood of collateral in a long-term loan by 2.5 percentage points, and by only 0.2 percentage points if the loan is short-term. Moreover, increases in the REAL INTEREST RATE tend to reduce the likelihood of collateral in short-term loans more than in long-term loans (0.4 and 0.3 percentage points, respectively). The same happens with the variable

GEOGRAPHIC RISK, which only has statistical and economic relevance for the short-term loans, a marginal effect of 0.2 percentage points.

An additional year of relations with the bank, DURATION, decreases the likelihood of collateral in short-term loans by 0.3 percentage points and almost four times more, 1.1 percentage points, in the long-term loans. A similar pattern is observed in the variable NUMBER OF LENDERS. A marginal increase in the number of loan operations with a bank significantly increases the likelihood of collateral in short-term loans more than in long-term loans (1.8 and 0.2 percentage points).

Although statistically significant, the variable of market concentration does not have economic significance to explain the use of collateral since its marginal effect is very low in all of the cases.

All of the marginal effects of *specialisation and preferences variables of the lender* are larger in absolute values in long-term loans than in short-term loans, while the marginal effect of the SIZE of the loan variable is quite small in the two groups of loans[15].

- *New borrowers*

Tables 5 and 6 present the results for new borrowers. The pattern of statistical and economic effects of the explanatory variables is quite similar to that observed in the group of old borrowers, so we shall limit our exposition to point out the main differences.

First, the proportion of loans with collateral among borrowers that default the loan one year after they get it is higher than in the group of borrowers that do not default, as was the case with old borrowers. But now the difference is half that obtained in the group of old borrowers (around 2.3 percentage points, marginal effects of the variable DEFAULTED (t+1), compared with approximately 4.5). Thus, the association between

---

[15] If the marginal effects were evaluated in relative terms vis-à-vis the predicted value of the dependent variable at the sample means (semielasticity), some of the conclusions from the comparison between the marginal changes in short- and long-term loans could change since the predicted probability of use of collateral is five times larger for long-term loans than for short-term loans (47.68% and 9.08%, respectively).

borrower risk and use of collateral is slighter in the group of new borrowers than in the group of old borrowers. Second, changes in the contemporaneous economic conditions, business cycle and interest rates, have a higher marginal effect in loans to new borrowers than in loans to old borrowers. Finally, among new borrowers marginal changes in the size of the loan increase the likelihood of collateral by 0.08 percentage points, but only by 0.01 percentage points among old borrowers.

4.2. Robustness Tests

One important limitation of the database used in this study is that the information about characteristics of the borrowers, such as size, debt ratios, quality of the management team, is not available. One way to control for these unobserved characteristics of the borrowers consists of estimating the probability models with borrowers' (firms) specific effects. In this way potential bias due to correlation between observed explanatory variables and unobserved ones will be corrected. At the same time the fixed effects mitigate the potential problem due to endogenous SIZE of the loan.

The econometric procedure used to introduce borrowers' fixed effects into the empirical model is described in Greene (1993). The procedure is based on a conditional likelihood approach implemented in a Logit model where it is possible to find a sufficient statistic for the fixed effect and after conditioning on it, the result obtained has similar properties to those obtained differentiating the data in conventional time and cross-section data panels. Therefore, the model to be estimated in this section is:

$$\Pr(Collateral_{it} = 1 \mid x_{i1},\ldots,x_{iT}, h_i) = \frac{\exp(x_{it} b + h_i)}{1+\exp(x_{it} b + h_i)},$$

where h_i denotes the borrower-specific effect.

The Logit model with firm-specific effects works in samples where each individual has at least a zero (the borrower gets a loan without collateral) and a one (the loan is with collateral). This substantially limits the sample of observations since only 25,199 (72,370) borrowers and 126,377 (357,527) loans satisfy this condition in the group of short- (long-) term loans; that is, the sample size in the group of old borrowers is one-fourth of the original one and the method can not be applied to new borrowers. Moreover, the new sample is biased towards larger firms, specially among the short-term loans, because the descriptive statistics (not reported) show that DEBT (t-1) and SIZE of the loan are up to five times larger in the restricted sample than in the population of loans. However, the Probit model with the restricted sample gives the same results (not reported) as determinants of the use of collateral as those shown in Table 4 for the whole population. The differences, if any, will have to be attributed to the control for fixed effects in each borrower.

The results from the Logit model with fixed effects are shown in Table 7. As could be expected, the comparison with the results in Table 4 shows that the main differences appear in the group of long-term loans and in the variables that capture attributes of the borrower such as DEBT (t-1), AGE, and relationship-lending. The negative sign of the coefficient of DEBT (t-1) would confirm that larger borrowers are less likely to pledge collateral in loans. Secondly, among long-term loans, the variables of relationship-lending all show coefficients with signs consistent with the hypothesis that relationship-lending lowers the likelihood of collateral (more duration, more loan operations and less number of lenders working will decrease the likelihood of collateral). Among short-term loans the evidence is less clear since, after controlling for the fixed effects, DURATION and SCOPE have coefficients with positive signs, and NUMBER of LENDERS has a coefficient with a negative sign.

The interpretation of the results from Tables 4 and 5 obtained for the whole population of loans will have to be made with some caution in the case of the relationship-lending hypothesis, but for the rest of the determinants of the use of collateral they appear quite robust.

4.3. Amount of collateral

The results of estimating the multinomial Logit model that explains the determinants of the amount of collateral in the population of old borrowers (i.e., no collateral at all, more than 50% of the loan, and 100% of the loan), are presented in Table 8. The model has the same explanatory variables as the model of the determinants of the use of collateral but the theoretical predictions are now limited to two of them, the REAL INTEREST RATE and the SIZE of the loan.

The first block of columns of Table 8 shows the coefficient (and statistics) of the model that explains the probability that a short-term loan will have a collateral of 100%. The second block of columns shows the result of the test of the difference between the coefficient of that model which explains the probability of 100% collateral and of the model which explains the use of collateral for at least 50% of the loan. A positive sign of this difference indicates that increases in the underlying explanatory variable increase more the likelihood of a loan with 100% collateral than the likelihood of a loan with partial collateral. The two blocks of columns are repeated for the group of long-term loans.

The difference between the two coefficients is positive, and statistically significant, for the variable REAL INTEREST RATE, and negative and statistically significant for the variable SIZE of the loan for both short-term and long-term loans. These results are consistent with the theoretical predictions.

As could be expected, the results of the model that explains that a loan will be of 100% collateral follow a pattern similar to those obtained in Table 4 for the model that explains the use of any amount of collateral[16]. The differences between the coefficients of the model that explains why a loan has 100% collateral, and the model that explains why a loan has only partial collateral, extend beyond the two variables indicated above. Notice that among those borrowers that default a loan in t-1 or in t+1, the likelihood of partial collateral is higher than the likelihood of total. This could indicate that more distressed borrowers have some difficulties to pledge 100 % collateral in their loans. On the other hand, the likelihood of 100% collateral increases more with the size of the loan than in the case of partial collateral.

5. Discussion

[16] The coefficients of the two models can not be directly compared since those of Table 4 are from a Probit model, while in Table 7 the model is a multinomial Logit one.

Credit markets are affected by severe information asymmetries between borrowers and lenders that give place to problems of adverse selection and moral hazard. To overcome these problems borrowers have incentives to signal their quality in terms of probability of defaulting the loan, and build long-term relationships with lenders in order to create a reputation of honest behaviour after receiving the loan. The lender, on the other hand, will research to know more about the credit risk of the borrower at the time the decision is made and may introduce debt covenants in the contract in order to curb the hidden-action problem. Theoretical and empirical models on the use of collateral in loan contracts have tried to explain how and why the presence of collateral in loans is consistent with the attempts to overcome problems of adverse selection and moral hazard.

The information contained in the Credit Register helps lenders to know the quality of the borrowers that ask for a loan, because it keeps updated information of the main characteristics of each loan, including information on whether the loan is in default or not. In this paper that information has been used to construct the variables DEFAULT (t-1) and DEFAULT (t+1). The first one indicates if, at the time of the loan decision, the lender knows if the borrower has a loan in default (a value of 1) or not (a value of 0). The empirical evidence that, with all being equal, a borrower with a 1 in the variable has an 86% (=7.81/9.08) higher probability of pledging collateral in a short-term loan than a borrower with a 0 value (31 per cent higher if the loan is long-term), provides strong evidence in favour of the prediction of Boot, Thakor and Udell (1991): if the lender knows the quality of the borrower, then low-quality borrowers will get a loan with collateral and high-quality borrowers will receive a loan without collateral.

One important finding of the paper is that the increase in the likelihood of the use of collateral when the variable DEFAULT (t+1) takes the value of 1, with respect to the likelihood when the value is 0, is smaller than the increase when the variable that changes from 0 to 1 is the variable DEFAULTED (t-1). Moreover, the change from 0 to 1 in the variable DEFAULTED (t+1) produces lower relative increases in the probability of collateral in the long-term loans than in the short-term ones (4.06/9.08 or 45% compared to 4.56/47.68 = 9.6%), and in the loans to new borrowers than in the loans to old borrowers (the absolute values of the changes are also lower in the former than in the latter according to Table 5). Therefore, even though the overall evidence is that the sorting according to observed risk dominates over the sorting according to private information, the relative importance of each of them differs under different circumstances also. First, sorting according to private information appears to be more frequent among borrowers who do not have a loan in default at the time of the loan decision than among those borrowers that have a loan in default at that time and are unambiguously classified as bad-quality borrowers. Second, sorting by private information is relatively more present among long-term loans than among short-term ones, which is consistent with the hypothesis that the information advantage of the borrower will increase with the maturity of the loan. And third, the increase in the probability of the use of collateral among those that default a loan in t+1 is lower among new borrowers, the group which has less experience in the credit market, and therefore lenders have had fewer opportunities to know about their quality.

An alternative way to evaluate the relationship between quality of the borrower and use of collateral is to compare the association between the use of collateral and credit risk in each of the situations considered. Previous work (Berger and Udell, 1990, and Jiménez and Saurina, 2003), find that the presence of collateral in a loan signals that the loan has higher credit risk than a loan with no collateral. From the results above, we can compute the conditional probability that the holder of the loan is going to be in default in the following time-period, given that the loan has collateral[17]. These probabilities are presented in Table 8 for new and old borrowers and for short- and long-term loans, together with the marginal probability of DEFAULT (t+1). The difference between the conditional and the marginal probability will be indicative of the information content of the signal "use of collateral in the loan".

The values shown in Table 9 indicate that the difference between the marginal and the conditional probability is higher in old borrowers than in new borrowers. For example, the conditional probability of default in t+1, given that the borrower is old (new) and the loan is short-term, is 15.43% (20.94%), a 36.6% (20.5%) higher than the marginal probability of default in the sample, 11.3% (17.37%). For long-term loans the information that the loan has collateral increases the probability of default in t+1, relative to the sample probability, by 7.5% in the case of old borrowers and 2.5% for the

[17] From the Bayes rule, Pr(Defaulted(t+1)/Collateral) = Pr(Collateral and Defaulted(t+1))/ Pr(Collateral), where Pr(Collateral and Defaulted(t+1)) = Pr(Collateral/Defaulted(t+1)) x Pr (Defaulted(t+1)).

new borrowers. Therefore, even though the use of collateral is a signal of a low-quality borrower, the information content of the signal is substantially higher in short-term loans than in long-term ones and among older borrowers than among new borrowers. As expected, collateral is a signal of risk in those populations of loans which are more frequently granted under the sorting by the observed-information paradigm than in populations where the dominant paradigm is sorting by private information.

The effects of reputation concerns of the borrower in the functioning of the loans market are captured by variables that measure the strength and scope of the relationship between borrowers and lenders (relationship-lending). To our knowledge this paper is the first one that provides empirical evidence on the effect of relationship-lending in the collateral decision controlling for firm-/borrower-specific effects. The control seems specially relevant since such an effect, as estimated by the econometric models, is quite different depending upon whether the fixed effects are introduced or not. For the whole population, and no fixed effects, we find that the evidence is mixed, since some variables, such as duration of the relationship, seem to indicate that reputation effects do in fact attenuate moral-hazard problems and lower the likelihood of collateral, as Boot and Thakor (1994) predict, but at the same time variables that mitigate the hold-up problem, such as NUMBER OF LENDERS, also reduce the likelihood of collateral. With fixed effects, in the sample of loans where it is possible to introduce them, the results are much clearer since they point, in a coherent way, towards the reputation hypothesis in the case of long-term loans, and more probably towards the hold-up one in the case of short-term loans. This evidence corroborates the point made by Berger and Udell (1995) that the investment in relationship-lending has different returns depending upon the type of loan[18]. It also points towards the recommendation to use panel data and control for borrower fixed effects in the study of credit decisions.

The paper provides evidence on the determinants of the use of collateral in loan contracts that have not been considered in previous studies, namely the impact of the business cycle and the efficiency and preferences of the lender. The macroeconomic conditions, phase of the business cycle and real interest rate of the economy affect the decisions on the use of collateral. In the first case the effect is asymmetric, in the sense

[18] We are unable to know how many of the financial loans are in fact lines of credit, the type of loans considered by Berger and Udell and by Harhoff and Körting in their analysis,

that lenders ask for collateral more frequently in the periods when the GDP grows at a rate slower than the long-term growth trend, but this effect is proportionally lesser than, and opposite to when GDP growth rates are above its average. The reaction to the macroeconomic conditions seems to be earlier and more intense, in absolute terms, in long-term loans than in short-term loans. The observed pattern is similar for old and new borrowers. On the other hand, the likelihood of collateral increases with the real interest rate of the economy, evidence that seems consistent with the hypothesis that in periods of tighter monetary conditions lenders give preference to their better customers in the allocation of credit.

Since credit markets are geographically segmented, the decision on the use of collateral may be affected by local market conditions such as competition among lenders and the state of the local economy. The geographic market considered in the paper is the province and the evidence shows that more competitive conditions in the local credit market, that is, a lower concentration of lenders, increases the likelihood of the use of collateral. This confirms the hypothesis that more competitive environments make relationship-lending more difficult and provide fewer economic incentives to invest in information about the risk and return of the loan operation; to ask for collateral may substitute for this investment (Manove and Padilla, 1999). The local-market economic conditions are evaluated through the percentage of bad loans over total loans in the province at the time of the loan and, as expected, the likelihood of the use of collateral increases with this ratio, i.e., lenders tend to ask for collateral more often in provinces where the local economic conditions are less buoyant. However, the economic significance of the credit-market concentration variable and of the geographic risk variable is very low, after controlling for idiosyncratic characteristics of these markets with the fifty provinces' dummy variables.

The experience and preferences of the lender may influence the decision to use collateral in a loan operation if more experience implies lower costs in the evaluation of the risk and return of the loan operation and if some lenders are willing to take more risks than others. We find that commercial banks have an economically significant lower probability of lending with collateral than savings banks (up to 25% lower, with respect to the predicted probability of collateral when the variables are at their sample means, in short-term loans and 12% lower in long-term loans). Since the former have more experience than the latter in the financial loan market to firms and, according to the descriptive data, tend to specialise in loans of larger size for which the fixed cost of evaluation can be spread in more units of principal, the efficiency advantage translates into less use of collateral. Besides, savings banks may be controlled by more risk-adverse managers than commercial banks. This evidence would confirm the relevance of variables that determine the expertise and resources of the lender to evaluate the loans operation, in deciding whether collateral will be a substitute of such evaluation, as Manove and Padilla (1999, 2001) claim is likely to happen.

One important novelty of this paper is that it presents evidence of the determinants of the amount of collateral controlling for the decision to pledge collateral. The results show that the amount of collateral will increase with the real interest rate of the economy and it will decrease with the size of the loan. Therefore, higher interest rates decrease the likelihood to use collateral in the loans made during that period, but once the decision is made to use collateral, its amount will increase with the interest rate. This evidence is consistent with the predictions of the model by Boot, Thakor and Udell (1991), which shows that, in order to avoid the negative effect on the effort of the borrower of higher interest rates in the presence of moral hazard, it is more efficient to ask for more collateral in the loan. Lenders ask for a lesser amount of collateral in loans of larger size because the payoff in the good states is also higher.

6. Conclusions

This paper investigates the determinants of the use of collateral in financial loans made to business firms, jointly with the determinants of the amount of collateral once the first decision is made. The research combines measures of *ex ante* and *ex post* credit risk to evaluate, in a novel and rigorous way, the association between credit quality of the borrower and the likelihood to use collateral, and it extends previous work in this topic to new variables that affect loan conditions, such as the business cycle, competition in the credit market and the expertise in evaluating loan operations and risk preferences of the lenders. The analysis takes advantage of a rich database, the Credit Register of the Bank of Spain, which allows us to construct a panel of data with the whole population of loans made to Spanish firms each year from 1984 to 2002. The empirical analysis also introduces some methodological novelties, such as to control for the possible endogenous nature of the SIZE variable of the loan, and to control for the fixed effects of the borrower (in this case in a sample from the whole population).

We find that low-quality borrowers, those who default a loan, are more likely to pledge collateral than high-quality ones. This is true when the borrower has defaulted a loan at the time of getting a new one and when the quality of the borrower is revealed *ex post*, after she receives a loan and defaults in the next period. This empirical evidence is consistent with other found in studies from other countries and consistent with the assumption that lenders can observe the quality of the borrower at the time of the loan decision. However, we also find that the association between quality of the borrower and the probability to pledge collateral is higher for short-term loans (maturity from one to three years) than for long-term loans, and that it is higher for older borrowers, those who have been in the loan market for at least one year, than for those that have been given a loan for the first time. These differences are interpreted as evidence that in those situations where the borrower is likely to have information advantage over the lender, when the maturity of the loan is longer and when the borrower is new, there are relatively more borrowers in the population which are sorted according to the private-information paradigm, i.e., a larger proportion of borrowers who consider themselves of high quality and who indicate their quality by choosing loans with collateral (and presumably lower interest rates). The observed probability of collateral conditioned to the quality of the borrower is the net result of borrowers sorted by observed information and borrowers sorted by private information.

The paper also finds that relationship-lending, longer relationships between borrower and lender, decreases the likelihood of collateral, as predicted by Boot and Thakor (1994), although the conclusion appears to be robust only in the sample of long-term loans. In the sample of short-term loans the conclusions are less clear and, if anything, they point towards the hypothesis of hold-up, (Sharpe, 1990). The results confirm that the returns from relationship-lending differ for each type of loan (Berger and Udell, 1995), and they strongly recommend the use of panel data in the study of credit decisions.

For the first time, to our knowledge, this paper provides evidence that general economic conditions affect the terms with which loans are made, in terms of both use and amount of collateral. When the economy is in a period of growth below the long-term trend, decreases in the GDP growth rate increase the likelihood that loans will be made with collateral. The effect of the business cycle in loan conditions seems to be stronger in loans to new borrowers, those that enter into the system for the first time, than in loans to established borrowers, while general economic conditions tend to trigger earlier responses in long-term loans than in short-term ones. The real interest rate of the economy affects the decision to use collateral in a different way than the decision over its amount, a negative effect in the first case and a positive one in the second. Lenders favour their preferred borrowers (of high quality) in times of tighter monetary conditions (higher real interest rates), but for those loans that will have collateral, the concern for a reduction in the effort of the borrower in response to higher interest rates moves the lender to the more efficient decision of raising the amount of collateral instead (Boot, Thakor and Udell, 1991).

There has been some concern in the credit market literature for the possibility that lenders have less incentive to invest time and resources in evaluating the risk of the loan operations due to the possibility of using collateral as a substitute for such evaluation (Manove and Padilla, 1999, 2001). Our results confirm that such may be the case since we find that those banks with less experience in the market of loans to business firms (savings banks) and those with less capabilities to evaluate loans (smaller banks) tend to ask more often for collateral in their financial loans than more experienced and more capable banks.

The findings of the paper have some relevant implications. Credit markets seem to function in a different way in loans of different maturity, short- or long-term, when the population of new borrowers who get a loan for the first time is considered, and when we look at lenders with more experience in the market. Research on the determinants of loan terms and conditions should focus on more segmented credit markets rather than considering the market as a whole. Large databases will be more suited for this new line of research, and in particular data that covers several time-periods, since evidence shows that the business cycle and other general economic conditions, such as the interest rate of the economy, play a role in determining some credit conditions such as the use and amount of collateral.

Banking regulators often look for signals that are predictors of risk and other sources of financial instability. The presence of collateral in loans can be one of these signals and, in fact, more use of collateral in loans will indicate that loans go to borrowers of low-quality, higher risk. Collateral is a way to curve down the risk of the loan operation taking into account the influence of the characteristics and behaviour of the borrower in such a risk. But, consistently with the general observation above, that the credit markets appear highly segmented, the information content of the signal use of collateral in the loan is substantially different among new borrowers than among old borrowers, and among short-term loans (with 10% of collateralised loans in our database) than among those long-term (where 50% of the loans have collateral).

Basel II will create incentives to lenders to invest in a more thorough analysis of the credit risk of borrowers. Evidence seems to indicate that lenders differ in their expertise and capabilities to evaluate the risk of the loans they granted, something that may not be so relevant in a time when the benefits of investing to gain such expertise and capabilities are low but that may be more important in the future under the new regulation of capital. One important research topic that has come out of our analysis is to investigate whether the new regulations in the capital requirements of banks may change the way banks view the use of collateral in loan operations.

References

Berger, A.N. and G. F. Udell (1990): "Collateral, Loan Quality, and Bank Risk." Journal of Monetary Economics, Vol. 25, pp. 21-4.

Berger, A.N. and G. F. Udell (1995): "Relationship Lending and Lines of Credit in Small Firm Finance." Journal of Business, Vol. 68, pp. 351-382.

Berger, A.N. and G. F. Udell (2002): "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organization Structure" Economic Journal, Vol. 112, pp 32-53.

Besanko, D. and A. V. Thakor (1987a): "Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets." International Economic Review, Vol. 28, pp. 671-689.

Besanko, D. and A. V. Thakor (1987b): "Competitive Equilibria in the Credit Market Under Asymmetric Information." Journal of Economic Theory, Vol. 42, pp. 167-182.

Bester, H. (1985): "Screening vs. Rationing in Credit Markets with Imperfect Information." American Economic Review, Vol. 75, pp. 850-855.

Boot, A.W.A. and A. V. Thakor (1994): "Moral Hazard and Secured Lending in an Infinitely Repeated Credit Market Game" International Economic Review, Vol. 35, No 4, November, pp. 899-92.

Boot, A.W.A., A. V. Thakor and G. F. Udell (1991): "Secured Lending and Default Risk: Equilibrium Analysis, Policy Implications and Empirical Results." The Economic Journal, 101, pp. 458-472.

Boot, A. W. A. (2000): "Relationship Banking: What Do We Know?" Journal of Financial Intermediation 9, pp. 7-25.

Carey, M.; M. Post and S. A. Sharpe (1998): "Does corporate lending by banks and finance companies differ? Evidence on specialization in private debt contracting." Journal of Finance, Vol LIII, no 3, pp. 845-878.

Chan, Y.S., S.I. Greenbaum and A.V. Thakor (1986):"Information reusability, competition and bank asset quality". Journal of Banking and Finance 10, pp. 255-276.

Chan, Y.S. and G. Kanatas (1985): "Asymmetric Valuation and the Role of Collateral in Loan Agreements." Journal of Money, Credit and Banking 17, pp. 85-95.

Degryse, H. and P. Van Cayseele (2000): "Relationship Lending within a Bank-Based System: Evidence from European Small Business Data" Journal of Financial Intermediation 9, pp. 90-109.

Diamond (1991): "Monitoring and reputation: The choice between bank loans and privately placed debt". Journal of Political Economy, Vol 99, pp. 689-721.

Gorton, G. and R. Rosen (1995): "Corporate Control, Portfolio Choice, and the Decline of Banking." The Journal of Finance, Vol. L, nº 5, December.

Greene, W. H. (1993): "Econometric Analysis." Macmillan.

Greenbaum, S., G. Kanatas and I. Venezia (1989): "Equilibrium loan pricing under the bank-client relationship". Journal of Banking and Finance, Vol 13, pp. 221-235.

Harhof, D. and T. Körting (1998): "Lending relationships in Germany-Empirical evidence from survey data. Journal of Banking and Finance, Vol 22, pp. 1317-1353.

Hester, D.D. (1979): "Customer Relationships and Terms of Loans." Journal of Money, Credit, and Banking, Vol. 11, pp. 349-357.

Jiménez,G. and J. Saurina (2003):"Collateral, Type of Lender and Relationship Banking as Determinants of Credit Risk". Forthcoming, Journal of Banking and Finance.

Manove, M. and A. J. Padilla (1999): "Banking (Conservatively) with Optimists." RAND Journal of Economics, Vol. 30, pp. 324-350.

Manove, M. and A. J. Padilla (2001): "Collateral versus Project Screening: a Model of Lazy Banks." RAND Journal of Economics, Vol. 32, nº 4, pp. 726-744.

Orgler, Y.E. (1970): "A Credit Scoring Model for Commercial Loans." Journal of Money, Credit, and Banking, Vol. 2, pp. 435-445.

Petersen, M.E. and R. G. Rajan (1994): "The Benefits of Firm-Creditor Relationships: Evidence from Small Business Data." The Journal of Finance 49, pp. 3-37.

Petersen, M.E. and R. G. Rajan (1995): "The Effect of Credit Market Competition on Lending Relationships." Quarterly Journal of Economics, Vol. 110, pp. 407-444.

Peek, J. and E. S. Rosengren (1996): "Small business credit availability: How important is size of lender? In "Financial system design: The case of universal banking" (A. Saunders and I. Walter, Eds.), pp. 628-655. Irwin, Burr Ridge.

Rajan, R.G. (1992): "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." The Journal of Finance 47, pp. 1367-1399.

Rajan, R.G. and A. Winton (1995): "Covenants and Collateral as Incentives to Monitor." The Journal of Finance, Vol. 50, pp. 1113-1146.

Rivers, D. and Q. H. Vuong (1988): "Limited Information Estimators and Exogeneity Test for Simultaneous Probit Models." Journal of Econometrics 39, pp. 347-366.

Saunders, A.; E. Strock and N. Travlos (1990): "Ownership Structure, Deregulation, and Bank Risk Taking." The Journal of Finance, Vol. XLV, nº 2, pp. 643-654.

Sharpe, S.A. (1990): "Asymmetric Information, Bank Lending, and Implicit Contracts: a Stylised Model of Customer Relationships. The Journal of Finance 45, pp. 1069-1087.

Table 1

Variables' Definitions

List of Variables Used in the Empirical Analysis and Their Respective Definition.

Variables	Definition
Borrower's Risk	
Defaulted $_{t-1}$	1 if the borrower has defaulted a loan anytime in previous year, 0 otherwise
Default $_{t+1}$	1 if the loan defaults in the first year but not in t-1, 0 ohtterwise
Debt $_{t-1}$	Total amount of debt of the borrower at the beginning of the period
"Age" as Borrower $_{t-1}$	Number of years since the first recorded loan to the borrower at the beginning of the period
Economic conditions	
GDP growth $_{t}$	Growth of the GDP in the year the loan is made
Phase economic cycle $_{t}$	Absolute difference between GDP growth in t and averaged GDP growth in t in the sample
GDP growth $_{t-1}$	Growth of the GDP in the previous year the loan is made
Phase economic cycle $_{t-1}$	Absolute difference between GDP growth in t-1and averaged GDP growth in t-1 in the sample
Real interest rate $_{t}$	Interest 3 months interbank minus rate of change in CPI in the year of the loan
Geographic risk $_{t-1}$	Proportion of bad loans in the province of the borrower at the beginning of the period
Product Market risk $_{t-1}$	Proportion of bad loans in the economic sector of the borrower at the beginning of the period
Relationship lending	
Duration $_{t-1}$	Number of years since the borrower had the first loan with the lender at the beginning of the period
Scope $_{t-1}$	Number of loan operations with the lender at the beginning of the period
Number of lenders $_{t-1}$	Number of lenders with which the borrower has relations at the beginning of the period
Competition	
Herfindhal $_{t}$	Sum of squared market shares of the lenders in the loan market of the province
Experience and preferences of lender	
Size of the lender $_{t}$	Assets of the lender relative to the total assets of the system
Bank	1 if the lender is a commercial bank, 0 otherwise
Specialisation $_{t}$	Percentage of loans to firms by the lender
Control Variables	
Size of the loan	Amount of the loan operation
Province	Dummy variable for the (50) provinces of geographic markets
Sector	Dummy variable for the (11) economic sectors of the lender
Dependent variables	
Collateral	1 if the loan has collateral, 0 otherwise
Amount of collateral	0 if the loan does not have collateral
	1if the loan is secured in more than 50 per cent but less than 100 per cent
	2 if the loan is 100 per cent secured

Table 2

Descriptive Statistics: Old Borrowers

Summary Statistics of the Variables Used in the Analysis for the Group of Old Borrowers. Time Period 1985 to 2002. For the definition of the variables see Table 1.

Variables	Mean		S.D.		Minimum		Maximum	
	Short term	Long term	Short term	Long term	Short term	Long term	Short term	Long term
Borrower's Risk								
Defaulted $_{t-1}$	0.03	0.04	0.16	0.20	0.00	0.00	1.00	1.00
Default $_{t+1}$	0.11	0.09	0.32	0.29	0.00	0.00	1.00	1.00
Debt $_{t-1}$	7.83	15.97	127.23	193.18	0.00	0.00	6234.27	6234.27
"Age" as Borrower $_{t-1}$	5.38	5.66	4.16	4.28	1.00	1.00	18.00	18.00
Economic conditions								
GDP growth $_t$ (%)	3.14	3.14	1.66	1.66	-3.41	-3.41	3.39	3.39
Phase economic cycle $_t$ (%)	1.88	1.88	1.16	1.16	0.25	0.25	3.41	3.41
Real interest rate $_t$ (%)	4.19	4.19	3.22	3.22	-1.01	-1.01	9.36	9.36
Geographic risk $_{t-1}$ (%)	5.77	5.77	4.32	4.32	0.48	0.48	28.34	28.34
Product Market risk $_{t-1}$ (%)	4.38	4.38	3.47	3.47	0.05	0.05	14.86	14.86
Relationship lending								
Duration $_{t-1}$	2.89	2.95	3.24	3.40	0.00	0.00	18.00	18.00
Scope $_{t-1}$	1.54	1.46	1.65	1.97	0.00	0.00	371.00	637.00
Number of lenders $_{t-1}$	3.93	3.94	6.69	9.71	0.00	0.00	184.00	184.00
Competition								
Herfindhal $_t$ (%)	6.48	6.87	2.64	2.67	2.52	2.52	42.48	42.48
Experience and preferences of lender								
Size of the lender $_t$ (%)	3.27	3.72	3.27	3.83	0.00	0.00	15.36	15.36
Bank	0.62	0.51	0.49	0.50	0.00	0.00	1.00	1.00
Specialisation $_t$ (%)	61.07	57.14	17.12	16.18	1.12	0.34	100.00	100.00
Control Variables								
Size of the loan	0.29	0.59	4.71	7.04	0.01	0.01	2937.16	2810.56
Dependent variable								
Collateral	0.08	0.53	0.28	0.50	0.00	0.00	1.00	1.00
% 100% collateralized	88.19	95.02						

Note: All the variables in levels. Debt and Size of the loan in millions of euros at constant prices of 2002.

Table 3

Descriptive Statistics: New Borrowers

Summary Statistics of the Variables Used in the Analysis for the Group of New Borrowers. Time Period 1985 to 2002. For the definition of the variables see Table 1.

Variables	Mean		S.D.		Minimum		Maximum	
	Short term	Long term	Short term	Long term	Short term	Long term	Short term	Long term
Borrower's Risk								
Default $_{t+1}$	0.17	0.09	0.38	0.29	0.00	0.00	1.00	1.00
Economic conditions								
GDP growth $_t$ (%)	3.14	3.14	1.66	1.66	-3.41	-3.41	3.39	3.39
Phase economic cycle $_t$ (%)	1.88	1.88	1.16	1.16	0.25	0.25	3.41	3.41
Real interest rate $_t$ (%)	4.19	4.19	3.22	3.22	-1.01	-1.01	9.36	9.36
Geographic risk $_{t-1}$ (%)	5.77	5.77	4.32	4.32	0.48	0.48	28.34	28.34
Product Market risk $_{t-1}$ (%)	4.38	4.38	3.47	3.47	0.05	0.05	14.86	14.86
Competition								
Herfindhal $_t$ (%)	6.36	6.75	2.63	2.70	2.52	2.52	42.48	42.48
Experience and preferences of lender								
Size of the lender $_t$ (%)	2.91	3.36	2.99	3.46	0.00	0.00	15.36	15.36
Bank	0.52	0.42	0.50	0.49	0.00	0.00	1.00	1.00
Specialisation $_t$ (%)	58.24	54.69	17.73	15.80	1.12	1.12	100.00	100.00
Control Variables								
Size of the loan	0.10	0.21	0.72	2.95	0.01	0.01	67.35	1341.85
Dependent variable								
Collateral	0.09	0.58	0.28	0.49	0.00	0.00	1.00	1.00
% 100% collateralized	90.93	96.47						

Note: All the variables in levels. Size of the loan in millions of euros at constant prices of 2002.

Table 4

Determinants of Collateral: Old Borrowers

Results of the Estimation of the Determinants of the Probability that a Loan in Year t Will Have a Collateral Given the Risk of the Borrower, the Macroeconomic Environment, the Concentration in the Geographic Market, the Relationship Between the Lender and the Borrower, the Characteristics of the Lender and the Size of the Loan as Control Variable, for the Group of Old Borrowers. Probit Model Estimated by 2SCML. For the definition of the variables see Table 1.

Variable	(1) Short term			(2) Long term		
Dependent Variable	Collateral (1/0)			Collateral (1/0)		
Estimation	2SCML			2SCML		
	Coefficient	*SD*	*p-value*	*Coefficient*	*SD*	*p-value*
Constant	-1.878	0.122	0.000	0.116	0.123	0.348
Borrower's Risk						
Defaulted $_{t-1}$	0.386	0.019	0.000	0.383	0.012	0.000
Default $_{t+1}$	0.225	0.012	0.000	0.117	0.011	0.000
log(Debt $_{t-1}$)	-0.073	0.046	0.114	0.018	0.024	0.465
log("Age" as Borrower $_{t-1}$)	-0.041	0.007	0.000	-0.010	0.004	0.009
Economic conditions						
GDP growth $_t$	-0.012	0.002	0.000	-0.011	0.002	0.000
Phase economic cycle $_t$	0.003	0.003	0.290	-0.014	0.003	0.000
GDP growth $_{t-1}$	-0.011	0.003	0.000	-0.063	0.002	0.000
Phase economic cycle $_{t-1}$	-0.023	0.003	0.000	-0.004	0.002	0.069
Real interest rate $_t$	-0.024	0.002	0.000	-0.008	0.001	0.000
Geographic risk $_{t-1}$	0.011	0.001	0.000	-0.001	0.001	0.386
Relationship lending						
log(Duration $_{t-1}$)	-0.046	0.007	0.000	-0.082	0.003	0.000
log(Scope $_{t-1}$)	0.171	0.014	0.000	0.009	0.006	0.140
log(Number of lenders $_{t-1}$)	-0.145	0.064	0.022	-0.304	0.036	0.000
Competition						
Herfindhal $_t$	-0.005	0.002	0.007	-0.008	0.001	0.000
Experience and preferences of lender						
Size of the lender $_t$	-0.013	0.001	0.000	-0.012	0.001	0.000
Bank	-0.107	0.022	0.000	-0.150	0.009	0.000
Specialisation $_t$	-0.005	0.001	0.000	-0.009	0.000	0.000
Control Variables						
log(Size of the loan)	0.384	0.080	0.000	0.192	0.045	0.000
Error log(Size of the loan)	-0.154	0.086	0.073	0.202	0.054	0.000
No. Observations	723,924			823,340		
% collateralized	8.48			54.24		
Log L	-187,870			-467,411		
χ^2 covariates	44,685			203,148		
Pseudo R^2	0.1063			0.1785		

Note: Each regression also includes 49 regional dummies and 10 industry dummies.

Table 5

Determinants of Collateral: New Borrowers

Results of the Estimation of the Determinants of the Probability that a Loan in Year t Will Have a Collateral Given the Risk of the Borrower, the Macroeconomic Environment, the Concentration in the Geographic Market, the Relationship Between the Lender and the Borrower, the Characteristics of the Lender and the Size of the Loan as Control Variable, for the Group of New Borrowers. Probit Model Estimated by 2SCML. For the definition of the variables see Table 1.

	(3)			(4)		
Variable	**Short term**			**Long term**		
Dependent Variable	Collateral (1/0)			Collateral (1/0)		
Estimation	2SCML			2SCML		
	Coefficient	*SD*	*p-value*	*Coefficient*	*SD*	*p-value*
Constant	-2.432	0.194	0.000	-1.363	0.214	0.000
Borrower's Risk						
Default $_{t+1}$	0.144	0.037	0.000	0.054	0.023	0.019
Economic conditions						
GDP growth $_t$	-0.025	0.004	0.000	-0.019	0.003	0.000
Phase economic cycle $_t$	-0.006	0.007	0.391	-0.015	0.004	0.001
GDP growth $_{t-1}$	-0.004	0.005	0.468	-0.045	0.003	0.000
Phase economic cycle $_{t-1}$	-0.027	0.006	0.000	0.001	0.004	0.758
Real interest rate $_t$	-0.036	0.004	0.000	-0.003	0.002	0.072
Geographic risk $_{t-1}$	0.009	0.003	0.001	-0.001	0.002	0.562
Competition						
Herfindhal $_t$	-0.007	0.004	0.049	0.000	0.002	0.947
Experience and preferences of lender						
Size of the lender $_t$	-0.010	0.002	0.000	-0.012	0.002	0.000
Bank	-0.135	0.026	0.000	-0.088	0.010	0.000
Specialisation $_t$	-0.003	0.001	0.022	-0.010	0.001	0.000
Control Variables						
log(Size of the loan)	0.433	0.067	0.000	0.459	0.052	0.000
Error log(Size of the loan)	-0.099	0.076	0.191	-0.007	0.063	0.915
No. Observations	170,317			254,755		
% collateralized	8.73			58.20		
Log L	-42,401			-136,833		
χ^2 covariates	16,129			72,612		
Pseudo R^2	0.1598			0.2097		

Note: Each regression also includes 49 regional dummies and 10 industry dummies.

Table 6

Marginal Effects

Marginal Changes in the Probability to Use Collateral in a Financial Loan in the Groups of Old and New Borrowers for Changes in the Explanatory Variables of Risk of the Borrower, Macroeconomic Conditions, Concentration of Lenders in the Geographic Market, Relationship Lending and Characteristics of the Lender. Estimations from the Probit Model of Tables 4 and 5. For the definition of the variables see Table 1.

	Old Borrowers		*New Borrowers*	
	Short term	**Long term**	**Short term**	**Long Term**
Predicted probability at sample mean (%)	9.08	47.68	8.84	50.60
	Marginal effect	Marginal effect	Marginal effect	Marginal effect
Borrower's Risk				
Defaulted $_{t-1}$	0.0781 ***	0.1478 ***	--	--
Default $_{t+1}$	0.0406 ***	0.0456 ***	0.0244 ***	0.0215 **
Debt $_{t-1}$	0.0000	0.0000	--	--
"Age" as Borrower $_{t-1}$	-0.0012 ***	-0.0007 ***	--	--
Economic conditions				
GDP growth $_t$	-0.0019 ***	-0.0043 ***	-0.0041 ***	-0.0078 ***
Phase economic cycle $_t$	0.0006	-0.0055 ***	-0.0009	-0.0058 ***
GDP growth $_{t-1}$	-0.0018 ***	-0.0247 ***	-0.0006	-0.0179 ***
Phase economic cycle $_{t-1}$	-0.0038 ***	-0.0015 *	-0.0043 ***	0.0005
Real interest rate $_t$	-0.0039 ***	-0.0030 ***	-0.0057 ***	-0.0014 *
Geographic risk $_{t-1}$	0.0018 ***	-0.0003	0.0014 ***	-0.0004
Relationship lending				
Duration $_{t-1}$	-0.0026 ***	-0.0108 ***	--	--
Scope $_{t-1}$	0.0179 ***	0.0024	--	--
Number of lenders $_{t-1}$	-0.0060 **	-0.0301 ***	--	--
Competition				
Herfindhal $_t$	-0.0008 ***	-0.0032 ***	-0.0011 **	0.0001
Experience and preferences of lender				
Size of the lender $_t$	-0.0021 ***	-0.0049 ***	-0.0016 ***	-0.0047 ***
Bank	-0.0231 ***	-0.0585 ***	-0.0216 ***	-0.0349 ***
Specialisation $_t$	-0.0008 ***	-0.0035 ***	-0.0005 **	-0.0038 ***
Control Variables				
Size of the loan	0.0002 ***	0.0001 ***	0.0007 ***	0.0009 ***

Notes: 1. All the parameters have been transformed in order to recover the original ones of the structural equation. The estimated probability at sample mean is computed as $\Phi(\hat{b}\overline{x})$ where $\overline{X}$ denotes the means of all the variables and $\hat{b}$ denotes the estimated parameters multiplied by $1\big/\sqrt{1+\hat{l}^2}$. The marginal effect is computed as $\frac{d\left[\Pr ob(y=1|x)\right]}{dx_k} = f(\hat{b}\overline{x})\hat{b}_k$ for all the variables except for debt, "age" as borrower, duration, scope, number of lenders and size of the loan which is $\left(f(\hat{b}\overline{x})\hat{b}_k\right)/\overline{x}_k$, to recover the effect of these variables in levels. For the binary ones is $\Pr(Collateral = 1|\overline{x}_*, d=1) - \Pr(Collateral = 1|\overline{x}_*, d=0)$, where $\overline{X}_*$ denotes the means of all the other variables. $\Phi(.)$ denotes the standard normal cumulative distribution function and $f(.)$ is its density function.

2. *** associated variable significant at the 1%, ** at the 5%, and * at the 10%.

Table 7

Determinants of Collateral: Old Borrowers with Fixed Effects

Results of the Estimation of the Determinants of the Probability that a Loan in Year t Will Have a Collateral Given the Risk of the Borrower, the Macroeconomic Environment, the Concentration in the Geographic Market, the Relationship Between the Lender and the Borrower, the Characteristics of the Lender and the Size of the Loan as Control Variable, for the Group of Old Borrowers. Logit Model with Borrowers' Fixed Effects. For the definition of the variables see Table 1.

	(5)			**(6)**		
Variable	**Short term**			**Long term**		
Dependent Variable	Collateral (1/0)			Collateral (1/0)		
Estimation	Logit with fixed effects			Logit with fixed effects		
	Coefficient	*SD*	*p-value*	*Coefficient*	*SD*	*p-value*
Borrower's Risk						
Defaulted $_{t-1}$	0.537	0.041	0.000	0.501	0.025	0.000
Default $_{t+1}$	0.326	0.030	0.000	0.220	0.020	0.000
log(Debt $_{t-1}$)	-0.021	0.010	0.035	-0.080	0.006	0.000
log("Age" as Borrower $_{t-1}$)	-0.012	0.017	0.455	0.002	0.010	0.825
Economic conditions						
GDP growth $_t$	-0.008	0.006	0.186	-0.004	0.004	0.338
Phase economic cycle $_t$	-0.001	0.010	0.909	-0.058	0.006	0.000
GDP growth $_{t-1}$	-0.042	0.008	0.000	-0.091	0.005	0.000
Phase economic cycle $_{t-1}$	-0.040	0.008	0.000	-0.020	0.005	0.000
Real interest rate $_t$	-0.051	0.005	0.000	-0.031	0.003	0.000
Geographic risk $_{t-1}$	0.026	0.003	0.000	0.012	0.002	0.000
Relationship lending						
log(Duration $_{t-1}$)	0.006	0.015	0.693	-0.039	0.008	0.000
log(Scope $_{t-1}$)	0.145	0.018	0.000	-0.097	0.011	0.000
log(Number of lenders $_{t-1}$)	-0.083	0.030	0.007	0.134	0.018	0.000
Competition						
Herfindhal $_t$	-0.014	0.007	0.027	-0.010	0.003	0.005
Experience and preferences of lender						
Size of the lender $_t$	-0.047	0.003	0.000	-0.038	0.001	0.000
Bank	-0.500	0.026	0.000	-0.480	0.015	0.000
Specialisation $_t$	-0.001	0.001	0.171	-0.010	0.000	0.000
Control Variables						
log(Size of the loan)	0.571	0.008	0.000	0.777	0.005	0.000
No. Observations	126,377			357,527		
% collateralized	29.67			50.25		
Log L	-39,687			-120,684		
χ^2 covariates	8,320			46,152		
Pseudo R^2	0.0949			0.1605		

Note: Each regression also includes 49 regional dummies and 10 industry dummies.

Table 8

Determinants of the Amount of Collateral: Old Borrowers

Results of the Estimation of the Determinants of the Amount of Collateral for a Loan in Year t Given the Risk of the Borrower, the Macroeconomic Environment, the Concentration in the Geographic Market, the Relationship Between the Lender and the Borrower, the Characteristics of the Lender and the Size of the Loan as Control Variable, for the Group of Old Borrowers. Estimation of a Multinomial Logic Model. For the definition of a variables see Table 1.

Variable	Short term (7)			(8)			Long term (9)			(10)		
Dependent Variable	Amount of collateral						Amount of collateral					
Estimation	2SCML						2SCML					
	100%			100% vs. 50%			100%			100% vs. 50%		
	Coefficient	*SD*	*p-value*	*Coefficient*	*SD*	*p-value*	*Coefficient*	*SD*	*p-value*	*Coefficient*	*SD*	*p-value*
Constant	-3.443	0.243	0.000	6.843	1.153	0.000	0.338	0.213	0.113	9.018	0.970	0.000
Borrower's Risk												
Defaulted $_{t-1}$	0.746	0.037	0.000	-0.346	0.166	0.037	0.624	0.021	0.000	-0.507	0.090	0.000
Default $_{t+1}$	0.450	0.024	0.000	-0.366	0.111	0.001	0.176	0.018	0.000	-0.565	0.082	0.000
log(Debt $_{t-1}$)	-0.134	0.091	0.141	1.507	0.433	0.000	0.124	0.042	0.003	1.115	0.193	0.000
log("Age" as Borrower $_{t-1}$)	-0.085	0.014	0.000	0.303	0.063	0.000	0.002	0.006	0.727	0.156	0.030	0.000
Economic conditions												
GDP growth $_t$	-0.022	0.004	0.000	0.067	0.019	0.000	-0.017	0.003	0.000	0.132	0.012	0.000
Phase economic cycle $_t$	0.003	0.007	0.664	-0.126	0.030	0.000	-0.010	0.004	0.027	-0.156	0.022	0.000
GDP growth $_{t-1}$	-0.032	0.005	0.000	-0.077	0.024	0.001	-0.111	0.003	0.000	-0.117	0.016	0.000
Phase economic cycle $_{t-1}$	-0.043	0.007	0.000	0.081	0.031	0.009	-0.006	0.004	0.089	0.081	0.017	0.000
Real interest rate $_t$	-0.035	0.005	0.000	0.113	0.023	0.000	-0.013	0.002	0.000	0.082	0.007	0.000
Geographic risk $_{t-1}$	0.020	0.002	0.000	-0.002	0.011	0.855	-0.003	0.001	0.015	-0.010	0.006	0.090
Relationship lending												
log(Duration $_{t-1}$)	-0.087	0.015	0.000	0.127	0.067	0.056	-0.145	0.005	0.000	0.001	0.024	0.950
log(Scope $_{t-1}$)	0.328	0.027	0.000	-0.626	0.126	0.000	-0.014	0.011	0.200	-0.599	0.049	0.000
log(Number of lenders $_{t-1}$)	-0.331	0.126	0.009	-2.358	0.600	0.000	-0.691	0.063	0.000	-1.762	0.288	0.000
Competition												
Herfindhal $_t$	-0.003	0.004	0.434	-0.056	0.016	0.000	-0.014	0.002	0.000	-0.045	0.010	0.000
Experience and preferences of lender												
Size of the lender $_t$	-0.007	0.002	0.001	0.172	0.012	0.000	-0.016	0.001	0.000	0.222	0.008	0.000
Bank	-0.082	0.043	0.057	-0.943	0.202	0.000	-0.253	0.015	0.000	-0.488	0.067	0.000
Specialisation $_t$	-0.013	0.002	0.000	0.029	0.009	0.001	-0.014	0.001	0.000	0.016	0.003	0.000
Control Variables												
log(Size of the loan)	0.719	0.160	0.000	-2.476	0.758	0.001	0.159	0.078	0.043	-2.069	0.358	0.000
Error log(Size of the loan)	-0.236	0.160	0.139	2.608	0.758	0.001	0.484	0.078	0.000	2.297	0.358	0.000
No. Observations	723,924						823,340					
% collateralized	7.48						50.57					
Log L	-210,025						-548,436					
χ^2 covariates	54,646						242,555					
Pseudo R^2	0.1151						0.1811					

Note: Each regression also includes 49 regional dummies and 10 industry dummies.

Table 9

Predicted Probabilities of Default in t+1

Conditional Probability that the Borrower will Default in Period t+1 Given that the Loan Granted in Period t-1 had Collateral. In Percentage Points

	Old Borrowers		*New Borrowers*	
	Short Term	**Long Term**	**Short Term**	**Long Term**
Average Probability of Defualt in t+1 for the population (from Tables 2 and 3)	11.30	9.10	17.37	9.12
Conditional Probability of Default in t+1 given that the Loan is Collateralized	15.43	9.78	20.94	9.35

Note: For example, the 15.43% is obtained as follows. The Pr(Collateral/Defaulted(t+1)) is obtained from the date with all other variables at their sample means and is equal to 12,73%. The marginal probability of default is Pr(Default(t+1))=11.30%. Therefore the joint probability of Collateral and Default will be Pr(Collateral and Default)= 1.40%. Since the marginal probability of Collateral is, for the variables at their sample means, Pr(Collateral)=9.08%, the conditional probability of Default given Collateral is equal to Pr(Default(t+1)/Collateral)=1.40/9.08=0.1543.

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Financial Intermediation 12 (2003) 25–56

Journal of Financial Intermediation

www.elsevier.com/locate/jfi

Estimating switching costs: the case of banking

Moshe Kim,[a,*] Doron Kliger,[a] and Bent Vale [b]

[a] *The University of Haifa, Haifa, Israel*
[b] *Norges Bank (Central Bank of Norway), Norway*

Received 3 October 2001

Abstract

We present an empirical model of firm behavior in the presence of switching costs. Customers' transition probabilities, embedded in firms' value maximization, are used in a multiperiod model to derive estimable equations of a first-order condition, market share (demand), and supply equations. The novelty of the model is in its ability to extract information on both the magnitude and significance of switching costs, as well as on customers' transition probabilities, from conventionally available highly aggregated data which do not contain customer-specific information. As a matter of illustration, the model is applied to a panel data of banks, to assess the switching costs in the market for bank loans. The point estimate of the average switching cost is 4.1%, about one-third of the market average interest rate on loans. More than a quarter of the customer's added value is attributed to the lock-in phenomenon generated by these switching costs. About a third of the average bank's market share is due to its established bank–borrower relationship.

JEL classification: L13; G21

Keywords: Switching costs; Transition probability; Banking

1. Introduction

Switching costs are costs induced to economic agents when they change their suppliers. As such, ex-ante homogeneous products become ex-post heterogeneous. These costs originate from a host of reasons, economic as well as psychological, such as various addictions and cognitive dissonance problems. Intertemporal product and service

* Corresponding author.
E-mail addresses: kim@econ.haifa.ac.il (M. Kim), kliger@econ.haifa.ac.il (D. Kliger), bent.vale@norges-bank.no (B. Vale).

1042-9573/03/$ – see front matter
doi:10.1016/S1042-9573(02)00005-0

compatibility, network externalities, informational investment in business relationships are a few examples of the economic sources of switching costs. From the theoretical perspective, customer switching costs confer market power on firms. Thus, firms may face a trade-off between charging low prices to attract customers and lock them in, and high prices to extract supranormal rents from its already locked-in customers. The vast theoretical literature on switching costs is summarized in Klemperer (1995).[1] Studies dealing with this phenomenon attempt to gain insight into the main issues of industrial organization such as entry deterrence, and the command over supranormal rents. Shapiro and Varian (1998) provide numerous examples of the impact of switching costs on market behavior.

The available empirical studies investigate the effect of switching costs on prices and market power. Ausubel (1991) provides some information that switching costs may explain the high interest rates on credit card balances, and Stango (1998), using variables related to switching, finds that switching costs have a significant impact on pricing in that market. Knittel (1997), using some proxies for switching costs, shows that they have provided long distance telephone carriers with market power. Sharpe (1997) finds that (banking) retail deposit-rates are positively affected by a proxy of switching costs. Dahlby and West (1986) support the effect of costly search on price dispersion in liability insurance, and Schlesinger and von der Schulenburg (1993) document a similar result in autoinsurance. Greenstein (1993) estimates the probability of "lock-in" in commercial mainframe computer systems acquired by federal agencies. His results may indirectly confirm the existence of switching costs for that sector but no quantification of the magnitude of switching costs is attempted. Another interesting empirical example is that of Borenstein (1991) for the gasoline market, where price discrimination is possible due to differences in the willingness of customers to switch stations. In a recent article, Shum (1999) measures the effect of advertising on habit persistence in the purchasing behavior of various brands of cereals. Shum finds that advertising encourages switching behavior at the household level. His main empirical question, however, concerns the way advertising affects brand substitutability thereby enhancing competitive conduct and lowering margins, and not the measurement

[1] Contributions to the theory of switching costs are as early as Selten (1965) and von Weizsäcker (1984). Klemperer (1985) examines a two-period differentiated-product duopoly in which customers are partially "locked-in" by switching costs that they face in the second period, which results in higher prices in both periods compared to the non switching cost case. Klemperer (1987) introduces switching costs in order to explain the emphasis placed on market share as a goal of corporate strategy. Beggs and Klemperer (1992) show how switching costs make the market more attractive to a new entrant and Klemperer (1987) examines how the threat of new entry affects incumbent's behavior thereby providing an explanation of limit pricing behavior. Chen and Rosenthal (1996) consider a stochastic game with slowly changing customer loyalties resulting in Markov perfect equilibrium. Padilla (1992) develops a model resulting in ex ante identical firms having ex post asymmetric market shares and in his (Padilla, 1995) paper he shows that in an infinite-horizon model with stationary Markovian strategies, the strength of competition is relaxed. Caminal and Matutes (1990) consider a model with endogenous switching costs and show that in the second period of the game firms discriminate against newcomers. Cabral and Greenstein (1990) claim that there may be economic merit to ignoring switching costs because the increased competitiveness in response to bidding parity can outweigh the costs of switching between suppliers.

of switching costs. Finally, a very recent paper by Israel (2001) develops and estimates a behavioral model of consumer–firm relationships in autoinsurance.[2]

Although the aforementioned empirical studies do point to the importance of switching costs in the determination of conduct and to the effect of various firms' policies on switching costs, they are generally silent regarding the magnitude and significance of switching costs. Whether switching costs are empirically important probably depends on the specific environment, industry, product type, and time period. One possible reason for the lack of empirical documentation of the magnitude and significance of switching costs is that the necessary micro data on individual-level transitions are rarely, if ever, available to researchers.[3] In the context of estimating switching costs, the unobservables are individual customers' purchase decision histories. More specifically, we lack information on the identity of customers' previous suppliers.

The task of the present research is to complement existing theoretical models with an empirical investigation capable of highlighting the process of customer's switching behavior when customer-specific data are absent, and then embed it in a general behavioral model of the firm. As a matter of illustration, we estimate the magnitude and significance of switching costs in the market for bank loans and empirically explore various counterfactuals related to bank and customer behavior in this market.[4]

It should be noted that from an empirical perspective, switching costs may be more pronounced when they contribute to, and may result from, long-term relationships and repeated contacts between firms and their customers. In such cases, customers' switching among suppliers may entail not only psychological costs but also costs related to the loss of capitalized value of established relationships. As such, the market for bank loans may provide a natural habitat for the investigation of the magnitude and significance of switching costs.

Banking is a major sector in the economy in which switching costs seem to be prevalent due to information asymmetry. A high quality borrower switching to a competing (uninformed) bank may be pooled with low quality borrowers and confront unfavorable conditions (Sharpe, 1990; von Thadden, 1998). This phenomenon (also known as the 'lemons' problem) may be exacerbated during periods of systemic wide banking problems or rescission periods. Thus, a switch between suppliers in the market for loans may entail direct (transaction-related) costs of closing an account with one bank and opening it elsewhere, as well as the unobserved, and perhaps the most significant costs associated with the foregone capitalized value of (previously established) long-term customer–bank

[2] Israel (2001) is interested in patterns of tenure-dependent demand whereas we are interested in estimating (the related phenomenon of) the magnitude of switching costs.

[3] There exist quite rich data sets on individual transactions pertaining to grocery-scanner, airline-reservation, and telephone-subscription data. However, even these data sets do not provide crucial information as to customer-specific purchase histories such as previous suppliers. Shum (1999), e.g., encounters the problem of having to use aggregate (national) data for his advertising expenditure which is the major variable in his model. Moreover, all these data sets are just (small) samples which may not represent the industry at large. This shortcoming may be crucial when specifying models of oligopolistic interaction among firms since the number of (relevant) firms in the industry (partially) determines pricing behavior.

[4] Our data set includes *all* banks in the industry, see Section 4.1.

relationship.[5] Indeed, the extensive discussion in recent literature about the importance of "relationship banking" and its significant impact on borrowers' values (James, 1987; Vale, 1993; Petersen and Rajan, 1994; Boot, 2000) may point to the existence of severe switching costs in this sector.[6]

The paper is organized as follows: in Section 2 we describe the model, and in Section 3 the empirical methodology. Section 4 describes the data used, and Section 5 provides the results and some counterfactuals. Section 6 concludes the paper.

2. A model of competition with switching costs

The empirical setup we present in this section builds on theoretical investigations related to the effect of customer switching costs on market conduct (see (Klemperer, 1987)). The theoretical relations between firm and customer behavior in the presence of switching costs require, however, some adjustment for empirical application. First, such relations are usually posited in terms of *rarely observable* individual switching activity. Secondly, while in most theoretical models switching costs serve as a threshold which, in equilibrium, causes customers *not to switch* between firms, casual empirical observation indicates to the contrary (see Klemperer (1995) for an overview).[7]

The model we propose in this study aims at providing behavioral equations fit for estimation, while remaining in line with, and in the spirit of, known theoretical results. We provide a setup which feeds on highly aggregated (conventionally available) data lacking information on customer-specific transition history. An additional attractive property of our model, vis-à-vis existing theoretical models, is that it allows customers to switch between firms at any period.

2.1. The model's framework

Consider an oligopoly of n firms competing in a multiple-stage price (Bertrand) competition. The good sold by the firms is unstorable. To focus on customers' decisions from which firm to purchase the good, the customers are assumed to have an inelastic demand.[8] Specifically, each customer purchases a single unit of the good at each one of

[5] Relationship- and transaction-related switching costs may reflect two extremes of the control-rent trade-off (Rajan, 1992). In the present model, however, we do not make a distinction between these due to lack of data on customers' transition histories.

[6] The importance of the relationship depends on the length of the interaction between the bank and its customer. Relationship may evolve in cases where complete contracts are not feasible, but long-term interaction is mutually beneficial. Thus, breaking up a relationship and forming a new one may entail severe switching costs. Also, the number of relationships may affect the value of relationships lost when switching suppliers. However, as found by Ongena and Smith (2000), the mean number of relationships is 1.37 and the median is 1.0. Thus, we expect quite a substantial informational loss due to switching.

[7] See von Thadden (1998) for a theoretical analysis showing switching of customers in equilibrium in the context of bank loans.

[8] This is often assumed in the literature (e.g., Caminal and Matutes (1990)).

an infinitely many discrete periods.[9] The customers maximize their utility by deciding from which firm to purchase, given the prices charged by the firms. When comparing the charged prices, the customers keep in mind that switching between firms is costly. The customer behavior described here yields probabilities of switching between firms. We dub these probabilities *transition probabilities*. The transition probabilities are functions of the prices and switching costs. Aggregation of the transition probabilities yields the demand faced by each firm.

2.2. Transition-probability induced demand

When choosing from which firm to purchase the good, the customer compares the prices charged by the various firms. We model the customer's purchase decision in terms of the probability of purchasing the good from the different firms. Such probabilistic modeling of the purchase decision allows customer heterogeneity in the model: when aggregating over the customers, the probability of purchasing from a specific firm proxies the proportion of customers who decide to buy from it. This formulation also facilitates the empirical observation that customers occasionally chose to switch to a firm charging a higher price. In fact, as long as the transition probabilities are not null or unity, some customers switch to and from each firm.

To accommodate the possible presence of switching costs, we condition the purchase probabilities on the customers' previous purchase decisions, i.e., we use *transition probabilities*. For tractability, we assume the transition probabilities are Markovian. Switching costs enter the model in the following way: for any possible purchase decision that entails switching (i.e., for purchasing from a firm the customer did not purchase from in the previous period), the cost of switching is added to the price charged by the firm. The probability that a customer who bought in period $t-1$ from firm i will retain her purchase in the subsequent period with that firm is denoted by $\Pr_{i\to i,t}$. Similarly, the probability that a customer who previously purchased from firm j will switch to purchase from firm i in the subsequent period is denoted by $\Pr_{j\to i,t}$. Formally, the transition probabilities are functions of the price charged by the firm, $p_{i,t}$, and of the alternative prices offered by firm i's rivals, the $(n-1)$ vector $\mathbf{p}_{iR,t}$.[10] The switching costs borne by the customer in case of switching to purchase from a firm she did not purchase from in the previous period are denoted by s.[11] The probability of continuing to purchase from the same firm is, therefore:

$$\Pr_{i\to i,t} = f\{p_{i,t}, \mathbf{p}_{iR,t} + \mathbf{s}\}, \tag{2.1}$$

where $\mathbf{s}$ is a vector of switching costs, equals the scalar s multiplied by an $(n-1)$ unity vector: $\mathbf{s} \equiv s \cdot I$.

[9] In the sequel, we relax this assumption and allow the demanded quantity to change at an exogenously determined rate.

[10] Throughout, we use bold letters to denote vectors.

[11] Actually, switching costs are likely to differ among customers. Therefore, s should be interpreted as the *mean* switching cost. Customer-specific deviations from within-firm averages are captured by the slope of the transition probability function, while the variation in the firm-mean of switching costs is captured by the level of the function.

The probability of switching to firm i from other firms is formulated as follows: If a customer of a rival firm j switches to *any* other firm (either to firm i, or to one of i's rivals), she bears the cost of switching. Thus, the (conditional) probability of switching from a rival firm j to firm i is:

$$\Pr_{j\to i,t} = f\{p_{i,t} + s, \mathbf{p}_{iR,t} + \mathbf{s}_j\}, \tag{2.2}$$

where $\mathbf{s}_j$ is an $(n-1)$ vector of switching costs in which each of the elements equals s, except for the jth element, which is zero. In aggregate data, transitions are not observed. We thus have to formulate the probability of switching to purchase from firm i *unconditional on the rival's identity*:

$$\Pr_{iR\to i,t} = \sum_{j\neq i}\left(f\{p_{i,t} + s, \mathbf{p}_{iR,t} + \mathbf{s}_j\}\cdot\frac{y_{j,t-1}}{\sum_{k\neq i} y_{k,t-1}}\right), \tag{2.3}$$

where $\Pr_{iR\to i,t}$ is the probability that a rivals' customer will switch to purchase from firm i. We denote firm j's time $t-1$ output by $y_{j,t-1}$. Thus, $y_{j,t-1}/\sum_{k\neq i} y_{k,t-1}$ is the probability that a *randomly selected* rival's customer is one who previously purchased from firm j.

A functional form for the transition probabilities should exhibit the following properties: the higher (lower) the relative price charged by firm i, the lower (higher) the probability that any customer will purchase from it. Thus, the partial derivatives of (2.1) and (2.3) should have the following signs:

$$\frac{\partial \Pr_{i\to i,t}}{\partial p_{i,t}} < 0, \qquad \frac{\partial \Pr_{i\to i,t}}{\partial \mathbf{p}_{iR,t}} > 0, \tag{2.4}$$

and

$$\frac{\partial \Pr_{iR\to i,t}}{\partial p_{i,t}} < 0, \qquad \frac{\partial \Pr_{iR\to i,t}}{\partial \mathbf{p}_{iR,t}} > 0. \tag{2.5}$$

The total demand faced by firm i at a time t, $y_{i,t}$, depends on its own output in the former period, as well as on the previous output of its rivals (the "state" variables), and on the transition probabilities which are functions of current prices and switching costs:

$$y_{i,t} = y_{i,t-1}\Pr_{i\to i,t} + y_{iR,t-1}\Pr_{iR\to i,t}. \tag{2.6}$$

The first term in the right-hand side of (2.6) approximates the number of the firm's customers that choose to continue purchasing from it. This is because applying the law of large numbers, $\Pr_{i\to i,t}$ proxies the *proportion* of firm i's "loyal" customers. Similarly, $\Pr_{iR\to i,t}$ is the proportion of new customers out of the rivals' customer base, making the second term in the right-hand side of (2.6) approximate the number of the rivals' customers that choose to switch to purchase from firm i.[12] Notice that $y_{i,t}$ is computed using information on firms' customer base, *without having to identify customer-specific actual* (*unobserved*) *transition decisions*.

[12] Since we would like to focus on customers' *transition probabilities* among firms, it is assumed that the number of customers is fixed. That is, we model switching behavior for any given level of aggregate demand in each period. Aggregate demand, however, is allowed to change over time.

To allow customers to change the purchased quantity over time, we multiply the demand faced by the firm by the market growth rate:

$$y_{i,t} = \Big(y_{i,t-1} \Pr_{i\to i,t} + y_{iR,t-1} \Pr_{iR\to i,t}\Big) g_t, \tag{2.7}$$

where g_t is the market growth rate, $g_t \equiv \sum y_{i,t} / \sum y_{i,t-1}$. We assume the market growth rate is exogenous.[13]

2.3. Approximating the demand

If *actual* customer decisions were observable, a natural structure to impose on the transition probabilities would have been of a logit type. However, aggregate data (a panel of market shares and prices) provide information on *net* output changes only. Any net output change could result from numerous combinations of customer arrivals and departures. Henceforth, we provide a setup in which the demand faced by each firm is a function of the firms' market shares. To obtain this setup, we apply a first-order (linear) approximation on the transition probabilities. Linearity makes the transition probabilities functions of the price charged by the firm, $p_{i,t}$, the average price charged by the rival firms, $\bar{p}_{iR,t}$, and the cost of switching between firms, s:[14]

$$\Pr_{i\to i,t} = \alpha_0^i + \alpha_1 p_{i,t} + \alpha_2(\bar{p}_{iR,t} + s), \tag{2.8}$$

and

$$\Pr_{iR\to i,t} = \alpha_0^i + \alpha_1(p_{i,t} + s) + \alpha_2\left(\bar{p}_{iR,t} + \frac{n-2}{n-1}s\right), \tag{2.9}$$

where α_0^i are firm-specific intercepts which capture firm heterogeneity.[15] The coefficient α_1 measures the sensitivity of the transition probability to the firm's own price, and α_2 is the cross-price sensitivity.[16]

It is important to emphasize here that (2.9) is not a function of a specific rival j, and thus, it is also the transition probability of a *randomly selected* rivals' customer. Note that the transition probabilities (2.8) and (2.9) remain valid for cases where the econometrician

---

13 Essentially, this assumption implies perfect foresight of the growth rate by banks, and it should be modeled in terms of expectations. For ease of notation, however, we omit the expectation operator. Note also that the growth rate in our data set is quite moderate (see Section 4.1).

14 While this specification is ad hoc, it maintains the major characteristics of a utility-driven demand: the probability of purchase (i) decreases with own price, and (ii) increases with other firms' prices.

15 The derivation of Eq. (2.9) is as follows. The transition probability from a rival firm j to firm i is $\Pr_{j\to i,t} = \alpha_0^i + \alpha_1(p_{i,t} + s) + \alpha_2^* I'(\mathbf{p}_{iR,t} + \mathbf{s}_j)$ where $\mathbf{s}_j$ is an $(n-1)$ vector of switching costs, in which each of the elements equals s, except for the jth element, which is zero. This can be written as $\Pr_{j\to i,t} = \alpha_0^i + \alpha_1(p_{i,t}+s) + \alpha_2^*(\sum_{j\neq i} p_{j,t} + (n-2)s)$, or, for $\alpha_2^* = \alpha_2/(n-1)$ $\Pr_{j\to i,t} = \alpha_0^i + \alpha_1(p_{i,t}+s) + \alpha_2(\bar{p}_{iR,t} + (n-2)s/(n-1))$. This equation is not a function of j, thus, it is also the transition probability of a *randomly selected* rivals' customer: $\Pr_{iR\to i,t} = \alpha_0^i + \alpha_1(p_{i,t}+s) + \alpha_2(\bar{p}_{iR,t} + (n-2)s/(n-1))$.

16 A more realistic, though untractable when applied to aggregate data, setup would allow the cross-price elasticities to vary among firms. Thus, the price sensitivities in our specification should be interpreted as average measures.

observes only a noisy version of the prices, such as prices which are unadjusted for output characteristics.[17]

As discussed in the previous section, the transition probabilities should be decreasing in the firm's own price and increasing in the average of the rivals' prices. Thus, the partial derivatives of (2.8) and (2.9), respectively, are expected to have the following signs:

$$\frac{\partial \Pr_{i\to i,t}}{\partial p_{i,t}} = \alpha_1 < 0, \qquad \frac{\partial \Pr_{i\to i,t}}{\partial \bar{p}_{iR,t}} = \alpha_2 > 0, \tag{2.10}$$

and

$$\frac{\partial \Pr_{iR\to i,t}}{\partial p_{i,t}} = \alpha_1 < 0, \qquad \frac{\partial \Pr_{iR\to i,t}}{\partial \bar{p}_{iR,t}} = \alpha_2 > 0. \tag{2.11}$$

Under the assumption of an inelastic total demand, a small increase in $p_{i,t}$ should have the same effect on the transition probabilities as that of a decrease of the same magnitude in rivals' average price, $\bar{p}_{iR,t}$. Thus, we restrict $\alpha_2 = -\alpha_1$. The resulting transition probabilities, then, are:

$$\Pr_{i\to i,t} = \alpha_0^i + \alpha_1(p_{i,t} - \bar{p}_{iR,t} - s), \tag{2.12}$$

and

$$\Pr_{iR\to i,t} = \alpha_0^i + \alpha_1\left(p_{i,t} - \bar{p}_{iR,t} - \frac{s}{n-1}\right). \tag{2.13}$$

As is apparent in Eqs. (2.12) and (2.13), the presence of switching costs works in favor, as well as against, the firm: the higher s, the higher the proportion of the firm's customers that will choose not to switch to other firms ($\partial \Pr_{i\to i,t}/\partial s = -\alpha_1 > 0$), but the lower the proportion of the rivals' customers that will choose not to switch to the firm ($\partial \Pr_{iR\to i,t}/\partial s = \alpha_1/(n-1) < 0$).

Using the linear transition probabilities, we are able to compute the firm's demand. Substituting (2.12) and (2.13) into the output equation (2.7) and dividing by the time t, market demand yields the firm's time t market share, $\sigma_{i,t}$:

$$\sigma_{i,t} = -\sigma_{i,t-1}\frac{n}{n-1}s\alpha_1 + \alpha_0^i + \alpha_1\left(p_{i,t} - \bar{p}_{iR,t} + \frac{s}{n-1}\right). \tag{2.14}$$

Thus, the period t market share is a function of the preceding period's market share, the number of competing firms, the switching costs, firms' heterogeneity (as captured by the fixed effects), and rivals' prices.

The part of the current market share induced by the time $t-1$ market share is represented by[18]

$$\left(-\sigma_{i,t-1}\frac{n}{n-1}s\alpha_1\right) > 0. \tag{2.15}$$

[17] The validity of the transition probabilities for this case is delegated to Appendix A.

[18] Recall that α_1 is negative.

We term the effect of the $t-1$ market share on the firm's current market share as the *lock-in effect*:

$$\frac{\partial\sigma_{i,t}}{\partial\sigma_{i,t-1}} = -\frac{n}{n-1}s\alpha_1 > 0. \tag{2.16}$$

The contribution of the existing (time $t-1$) market share to the current market share is increasing in the magnitude of switching-costs,

$$\frac{\partial(\partial\sigma_{i,t}/\partial\sigma_{i,t-1})}{\partial s} = -\frac{n}{n-1}\alpha_1 > 0. \tag{2.17}$$

We dub the effect of switching costs on market shares the *switching-cost effect*:

$$\frac{\partial\sigma_{i,t}}{\partial s} = \left(\frac{1}{n} - \sigma_{i,t-1}\right)\frac{n}{n-1}\alpha_1 \begin{cases} < 0 & \text{if } \sigma_{i,t-1} < 1/n, \\ > 0 & \text{if } \sigma_{i,t-1} > 1/n. \end{cases} \tag{2.18}$$

The switching costs effect works in favor of larger-than-average firms, and against smaller-than-average firms. The intuition behind this effect is straightforward. The larger the firm's market share, the more customers are "locked-in" with it, and the less customers are "locked out" of the firm.

2.4. Firms' present-value maximization

At any point in time (denoted by τ), the firm maximizes the present value of its profits:

$$V_{i,\tau} = \sum_{t=\tau}^{\infty} \delta^{t-\tau}\pi_{i,t}, \tag{2.19}$$

where δ is the one-period discount factor, $\pi_{i,t}$ is the firm's time t profit, i.e.,

$$\pi_{i,t} \equiv y_{i,t} \cdot p_{i,t} - c_{i,t}, \tag{2.20}$$

and the firm's technology is specified by a cost function, defined on its output, $y_{i,t}$, and on a vector of input prices, $\mathbf{w}_{i,t}$:

$$c_{i,t} = c\{\mathbf{w}_{i,t}, y_{i,t}\}. \tag{2.21}$$

A necessary condition for the firm's optimal (present value maximizing) behavior is setting the time τ price, $p_{i,\tau}$, such that the derivative of (2.19) w.r.t. this price is zero:[19]

$$\frac{\partial V_{i,\tau}}{\partial p_{i,\tau}} = \sum_{t=\tau}^{\infty} \delta^{t-\tau}\frac{\partial\pi_{i,t}}{\partial p_{i,\tau}} = 0. \tag{2.22}$$

Notice that the time τ price affects not only the time τ profit, $\pi_{i,\tau}$, but also the profits in subsequent periods. The reason for this is that any period's output affects the demand in the period that follows.[20]

[19] The second-order condition for present value maximization is satisfied because (2.19) is concave: the present value function is a sum of the concave per-period profits, $\pi_{i,t}$, and all future demanded quantities ($y_{i,\tau+k}$, $k > 1, 2, \ldots$) may be expressed as linear functions of the current demand, $y_{i,t}$.

[20] This is easily seen by iterative substitution of Eq. (2.7) which describes the demand faced by the firm.

Inserting the time t profit, (2.20), (2.22) becomes:

$$\frac{\partial V_{i,\tau}}{\partial p_{i,\tau}} = \sum_{t=\tau}^{\infty} \delta^{t-\tau} \frac{\partial (y_{i,t} \cdot p_{i,t} - c_{i,t})}{\partial p_{i,\tau}}$$
$$= y_{i,\tau} + \sum_{t=\tau}^{\infty} \delta^{t-\tau} \left(\frac{\partial y_{i,t}}{\partial p_{i,\tau}} p_{i,t} - \frac{\partial c_{i,t}}{\partial y_{i,t}} \frac{\partial y_{i,t}}{\partial p_{i,\tau}} \right) = 0 \tag{2.23}$$

or

$$\frac{\partial V_{i,\tau}}{\partial p_{i,\tau}} = y_{i,\tau} + \sum_{t=\tau}^{\infty} \delta^{t-\tau} \left(p_{i,t} - \frac{\partial c_{i,t}}{\partial y_{i,t}} \right) \frac{\partial y_{i,t}}{\partial p_{i,\tau}} = 0, \tag{2.24}$$

where the effect of the current price on the quantity demanded k periods ahead is:

$$\frac{\partial y_{i,\tau+k}}{\partial p_{i,\tau}} = \frac{\partial y_{i,\tau+k}}{\partial y_{i,\tau+k-1}} \cdot \frac{\partial y_{i,\tau+k-1}}{\partial y_{i,\tau+k-2}} \cdot \cdots \cdot \frac{\partial y_{i,\tau}}{\partial p_{i,\tau}} \quad \text{for } k = t - \tau. \tag{2.25}$$

For similar arguments, another requirement for the firm's optimal behavior is that the derivative of (2.19) w.r.t. the time $\tau + 1$ price, $p_{i,\tau+1}$, is zero along the optimal path:

$$\frac{\partial V_{i,\tau}}{\partial p_{i,\tau+1}} = y_{i,\tau+1} + \sum_{t=\tau}^{\infty} \delta^{t-\tau+1} \left(p_{i,t} - \frac{\partial c_{i,t}}{\partial y_{i,t}} \right) \frac{\partial y_{i,t}}{\partial p_{i,\tau+1}} = 0, \tag{2.26}$$

where

$$\frac{\partial y_{i,\tau+k}}{\partial p_{i,\tau+1}} = \frac{\partial y_{i,\tau+k}}{\partial y_{i,\tau+k-1}} \cdot \frac{\partial y_{i,\tau+k-1}}{\partial y_{i,\tau+k-2}} \cdot \cdots \cdot \frac{\partial y_{i,\tau+1}}{\partial p_{i,\tau+1}} \quad \text{for } k = t - \tau. \tag{2.27}$$

Since both (2.24) and (2.26) are necessary conditions, any linear combination of them should hold as well. Thus, for any $\mathrm{d}p_{i,\tau}$ and $\mathrm{d}p_{i,\tau+1}$, the following should hold:

$$\frac{\partial V_{i,\tau}}{\partial p_{i,\tau}} \mathrm{d}p_{i,\tau} + \frac{\partial V_{i,\tau}}{\partial p_{i,\tau+1}} \mathrm{d}p_{i,\tau+1} = 0. \tag{2.28}$$

In particular, let us choose a pair of price differentials $\mathrm{d}p_{i,\tau}$ and $\mathrm{d}p_{i,\tau+1}$ that keeps $y_{i,\tau+1}$ constant:

$$\frac{\partial y_{i,\tau+1}}{\partial p_{i,\tau}} \mathrm{d}p_{i,\tau} + \frac{\partial y_{i,\tau+1}}{\partial p_{i,\tau+1}} \mathrm{d}p_{i,\tau+1} = 0 \tag{2.29}$$

or

$$\mathrm{d}p_{i,\tau+1} = -\frac{\partial y_{i,\tau+1}}{\partial p_{i,\tau}} \Big/ \frac{\partial y_{i,\tau+1}}{\partial p_{i,\tau+1}} \mathrm{d}p_{i,\tau}. \tag{2.30}$$

Substituting

$$\partial y_{i,\tau+1}/\partial p_{i,\tau} = -y_{\tau-1}\alpha_1 \frac{n}{n-1} \alpha_1 s g_\tau g_{\tau+1} \quad \text{and}$$
$$\partial y_{i,\tau+1}/\partial p_{i,\tau+1} = y_{\tau-1}\alpha_1 g_\tau g_{\tau+1}$$

into (2.30), we obtain the time $\tau + 1$ price-differential[21] in terms of the time τ price-differential, for a constant $y_{i,\tau+1}$:

$$\mathrm{d}p_{i,\tau+1} = \mathrm{d}p_{i,\tau} \frac{n}{n-1} \alpha_1 s. \tag{2.31}$$

Since $y_{i,\tau+1}$ is unchanged, condition (2.28) becomes:[22]

$$\left(\frac{\partial \pi_{i,\tau}}{\partial p_{i,\tau+1}} + \delta \frac{\partial \pi_{i,\tau}}{\partial p_{i,\tau}} \right) \mathrm{d}p_{i,\tau} + \delta \frac{\partial \pi_{i,\tau+1}}{\partial p_{i,\tau+1}} \mathrm{d}p_{i,\tau+1} = 0. \tag{2.32}$$

Furthermore, as $y_{i,\tau+1}$ is constant, (2.32) becomes:

$$\frac{\partial \pi_{i,\tau}}{\partial p_{i,\tau}} \mathrm{d}p_{i,\tau} + \delta \cdot y_{i,\tau+1} \mathrm{d}p_{i,\tau+1} = 0. \tag{2.33}$$

Inserting (2.31) into (2.33) and rearranging:

$$\frac{\partial \pi_{i,\tau}}{\partial p_{i,\tau}} + \delta \cdot y_{i,\tau+1} \frac{n}{n-1} \alpha_1 s = 0. \tag{2.34}$$

Writing the derivative of the time τ profit explicitly, yields

$$y_{i,\tau} + \left(p_{i,\tau} - \frac{\partial c_{i,\tau}}{\partial y_{i,\tau}} \right) \frac{\partial y_{i,\tau}}{\partial p_{i,\tau}} + \delta y_{i,\tau+1} \frac{n}{n-1} \alpha_1 s = 0. \tag{2.35}$$

As $\partial y_{i,\tau} / \partial p_{i,\tau} = y_{\tau-1} \alpha_1 g_\tau$, (2.35) can be expressed as

$$pcm_{i,t} = -\delta \cdot \sigma_{i,t+1} \frac{n}{n-1} s g_{t+1} - \frac{\sigma_{i,t}}{\alpha_1}, \tag{2.36}$$

where $pcm_{i,t} \equiv p_{i,t} - mc_{i,t}$ is the period t price–cost margin. Equation (2.36) captures the relation between the price–cost margin, the market shares, and the switching costs. To provide some intuition for this first-order condition, notice that the first term represents the benefits to the firm from capturing customers in period t that will be “locked in” in future periods. The larger this benefit is (a higher s or g_{t+1}), the lower will be the optimal period t price–cost margin in the attempt to capture customers. The second term represents the current period oligopoly power of the firm; the larger is the current market share, the larger will be the price–cost margin. Notice that the existence of switching costs results in a larger market shares and hence more oligopoly power for larger than average firms (cf. Eq. (2.18)). In the absence of switching costs the optimization problem reduces to the conventional case of a one-period oligopoly,

$$pcm_{i,t} = -\frac{\sigma_{i,t}}{\alpha_1}, \tag{2.37}$$

where as far as switching costs exist, the firms’ pricing decision is intrinsically intertemporal.

An increase in the discount factor δ (i.e., a decrease in the discount rate) has the same effect on prices as an increase in the growth rate of aggregate output g_{t+1}.[23] A higher

[21] For a derivation of the output derivatives refer to Appendix C.

[22] The terms in (2.28) containing the derivatives of $\tau_{i,\tau+2}$ and onwards disappear as they all are, ultimately, products of the change in $y_{i,\tau+1}$, which is zero.

[23] This property is in line with the prediction of Beggs and Klemperer (1992).

growth rate increases the relative importance of future profits, hence making it more profitable for firms to lower current prices and lock in a larger market share.[24]

The firm's optimization may shed some light on the value of customer lock-in. Define the marginal value of a locked-in customer, $mvl_{i,t}$, as the marginal increase in the firm's present value due to an additional locked-in customer $(\partial V_{i,t}/\partial y_{i,t})$, beyond the increase in profits generated by the current sales to that customer. This implies that

$$\frac{\partial V_{i,t}}{\partial y_{i,t}} = \frac{\partial \pi_{i,t}}{\partial y_{i,t}} + mvl_{i,t}. \tag{2.38}$$

For constant market size in steady state, the proportion of $mvl_{i,t}$ in $\partial V_{i,t}/\partial y_{i,t}$ equals the lock-in effect (defined in (2.16)), discounted for one period:[25]

$$\frac{mvl_{i,t}}{\partial V_{i,t}/\partial y_{i,t}} = -\delta \frac{n}{n-1} s\alpha_1. \tag{2.39}$$

This is so because $mvl_{i,t}$ equals the discounted marginal increase in the firm's present value due to an additional locked-in customer in the subsequent period:

$$mvl_{i,t} = \delta \frac{\partial V_{i,t+1}}{\partial y_{i,t}} = \delta \frac{\partial V_{i,t+1}}{\partial y_{i,t+1}} \frac{\partial y_{i,t+1}}{\partial y_{i,t}}, \tag{2.40}$$

implying that

$$\frac{\partial V_{i,t}}{\partial y_{i,t}} = \frac{\partial \pi_{i,t}}{\partial y_{i,t}} + \delta \frac{\partial V_{i,t+1}}{\partial y_{i,t}}. \tag{2.41}$$

In this case, $\partial V_{i,t+1}/\partial y_{i,t+1} \approx \partial V_{i,t}/\partial y_{i,t}$, since (for $g_t = 1$, $\forall t$) $y_{i,t+1} \approx y_{i,t}$. Solving for $\partial V_{i,t}/\partial y_{i,t}$ yields:

$$\frac{\partial V_{i,t}}{\partial y_{i,t}} = \frac{\partial \pi_{i,t}}{\partial y_{i,t}} \left(1 - \delta \frac{n}{n-1} s\alpha_1\right)^{-1} \tag{2.42}$$

or, as a proportion of the added value of an additional customer,

$$\frac{\partial \pi_{i,t}/\partial y_{i,t}}{\partial V_{i,t}/\partial y_{i,t}} = \left(1 - \delta \frac{n}{n-1} s\alpha_1\right). \tag{2.43}$$

Using (2.38), the contribution of lock-in as a proportion of the added value is

$$\frac{mvl_{i,t}}{\partial V_{i,t}/\partial y_{i,t}} = 1 - \frac{\partial \pi_{i,t}/\partial y_{i,t}}{\partial V_{i,t}/\partial y_{i,t}} = -\delta \frac{n}{n-1} s\alpha_1. \tag{2.44}$$

3. Empirical methodology

The model presented in Section 2 yields the following equations:

[24] Slade (1991) provides empirical support for that effect for the metals markets.
[25] We believe constant market size is a reasonable approximation. See Section 4.

(i) the first-order condition (2.36):

$$pcm_{i,t} = -\delta\sigma_{i,t+1}\frac{n}{n-1}sg_{t+1} - \frac{\sigma_{i,t}}{\alpha_1}, \tag{3.1}$$

(ii) the market share Eq. (2.14):

$$\sigma_{i,t} = -\sigma_{i,t-1}\frac{n}{n-1}s\alpha_1 + \alpha_0^i + \alpha_1\left(p_{i,t} - \bar{p}_{iR,t} + \frac{s}{n-1}\right). \tag{3.2}$$

In order to estimate the model we search for the parameter values that best fit the data.[26] To eliminate the numerous fixed effects (α_0^i), the market share Eq. (3.2) is first-differenced. To obtain the price–cost margin in (3.1), we estimate the marginal cost implied by the system of cost function and input share equations (see Appendix B).[27] The system of equations (including the cost structure) is estimated simultaneously using non-linear 3SLS. The endogenous variables in the model $\sigma_{i,t}$, $\sigma_{i,t+1}$, the time differences of the prices and of the market shares as well as the output, are instrumented by the lagged output, lagged number of branches and various lags of the market shares.

A positive s is our basic indication for the existence of customer switching costs in the market for bank loans. We interpret the data as providing significant evidence for the existence of switching costs if s turns out significantly positive. Furthermore, since demand should be downward sloping, the validity of the model is indicated by the sign of α_1. The demand curve is downward sloping if and only if α_1 is negative.

3.1. Details

In the estimation procedure, we have to make sure the transition probabilities are within the $[0, 1]$ interval. First, we show here the range of switching costs for which the linear approximation is valid. Then, we show how to impose this range on the estimated switching costs.

To derive the valid range for the switching costs parameter, recall the formulas of the transition probabilities, (2.8) and (2.9):

$$\Pr_{i\to i,t} = \alpha_0^i + \alpha_1(p_{i,t} - \bar{p}_{iR,t} - s) \tag{3.3}$$

and

$$\Pr_{iR\to i,t} = \alpha_0^i + \alpha_1\left(p_{i,t} - \bar{p}_{iR,t} + \frac{s}{n-1}\right). \tag{3.4}$$

When switching costs are high, (3.3) approaches unity and (3.4) approaches zero, i.e., most of the customers are effectively locked-in with their former suppliers. For the linear transition probabilities to be valid, they have to lie within the $[0, 1]$ interval. The following

[26] We use a panel data of the entire Norwegian banking sector for the period 1988–1996. For details see Section 4.

[27] We note that in the empirical application we have also experimented with a short run variable cost function where equity was treated as a quasi fixed factor, as suggested by Hughes and Mester (1998). Results where almost identical to the those reported here.

should, thus, hold:

$$-1 \leqslant \Pr_{i\to i,t} - \Pr_{iR\to i,t} = -\alpha_1 s \frac{n}{n-1} \leqslant 1 \tag{3.5}$$

or

$$\frac{1}{\alpha_1}\frac{n-1}{n} \leqslant s \leqslant \frac{1}{-\alpha_1}\frac{n-1}{n}. \tag{3.6}$$

Note that if, in the estimation process, s is outside the range depicted in (3.6), the estimated equations become meaningless, since it implies negative or larger-than-unity transition probabilities.[28] To avoid the introduction of inequality constraints into the estimated system, define:

$$\bar{s} \equiv \frac{1}{-\alpha_1}\frac{n-1}{n} \tag{3.7}$$

and, instead of directly estimating s, estimate a transformation of it, $z \in \Re$ defined by

$$s \equiv 2\bar{s} \cdot \frac{e^z}{1+e^z} - \bar{s}. \tag{3.8}$$

Note that as $z \to -\infty$, $s \to -\bar{s}$, and as $z \to \infty$, $s \to \bar{s}$. In the estimated equations, we replace s by $2\bar{s} \cdot e^z/(1+e^z) - \bar{s}$ everywhere.[29]

4. Data and industry characteristics

4.1. The banking industry

Our database consists of a panel of annual observations for the Norwegian banking industry, spanning nine years from 1988 to 1996. The panel covers *all* banks operating in Norway in that period.[30] Table 1 describes the banking industry characteristics.

The number of banks declined from 177 in 1988 to 139 in 1996. The reduction in the number of banks is almost only due to mergers.[31] There is a high cross-sectional variation in bank size, most banks being very small. Figures 1 and 2 plot the observations by number of branches, and Figs. 3 and 4 by market shares, measured by value of loans.

---

[28] Actually, n changes slightly over time. Thus, for the constraint to hold in all time periods, the minimal n is applied. In any case, applying different n values is of minor effect, due to the minor variation in n, and since $\partial \frac{n-1}{n}/\partial n \to 0$ for large n (there were 139–177 banks in Norway during that time).

[29] As turned out in actual estimation, s was inside the range depicted in (3.6). Thus, the transformation of s to z was not needed, and a direct estimation of s was carried out.

[30] The Post bank and several state banks were excluded due to their different nature of business and problematic data.

[31] Most of the mergers are characterized by one relatively large bank (the predator) buying a smaller bank (the prey). Since in these cases the switching costs may be lower or zero, we choose to mark customers of the prey as "switchers," thereby probably downward biasing the switching cost estimate. An opposite, upward biasing, treatment would be to exclude all banks that have experienced mergers. Unfortunately, the estimated equation system did not converge when applying this treatment to our data set.

Table 1
Industry characteristics

	1988	1989	1990	1991	1992	1993	1994	1995	1996
Number of banks	177	168	155	147	145	143	141	140	139
No. of branches:									
– mean	12	12	12	11	11	11	11	11	11
– max	160	141	240	205	193	182	182	194	183
Bank Loans (bil. NOK)	423	456	473	444	443	453	471	512	580
Loans per bank (bil. NOK):									
– mean	2.38	2.71	3.05	3.02	3.06	3.17	3.34	3.66	4.17
– std. dev.	8.16	9.46	13.44	12.50	12.65	11.92	11.98	12.94	14.82
Equity per bank (bil. NOK):									
– mean	0.13	0.15	0.15	0.11	0.13	0.23	0.29	0.34	0.38
– std. dev.	0.45	0.56	0.61	0.28	0.23	0.66	1.01	1.17	1.34
Assets per bank (bil. NOK):									
– mean	3.24	3.56	3.93	3.86	4.00	4.02	4.19	4.44	5.21
– std. dev.	11.43	12.60	17.49	16.15	17.17	15.69	15.74	16.14	19.48
Interest rate (pct.):									
– mean	16.07	14.82	14.15	13.72	13.23	10.99	8.47	8.11	7.50
– std. dev.	1.32	0.89	0.79	0.69	1.38	0.69	0.57	0.48	0.48
T-bill rate (pct.)	12.88	10.86	10.86	9.99	9.62	6.86	7.46	7.43	6.78

Bank loans include all domestic loans extended by all banks in the sample. Interest rate is the calculated interest rate on loans extended by the banks in the sample. See Section 4.2 for calculation details.

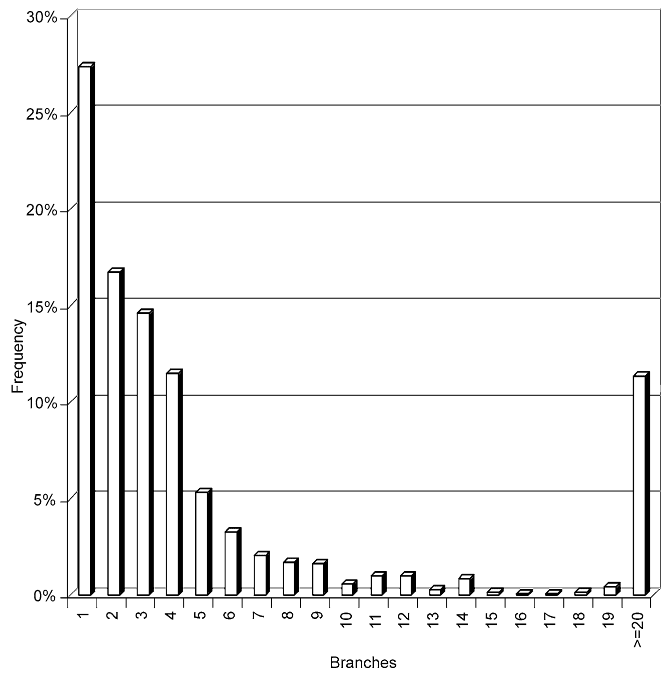


Fig. 1. Sample distribution by number of branches.

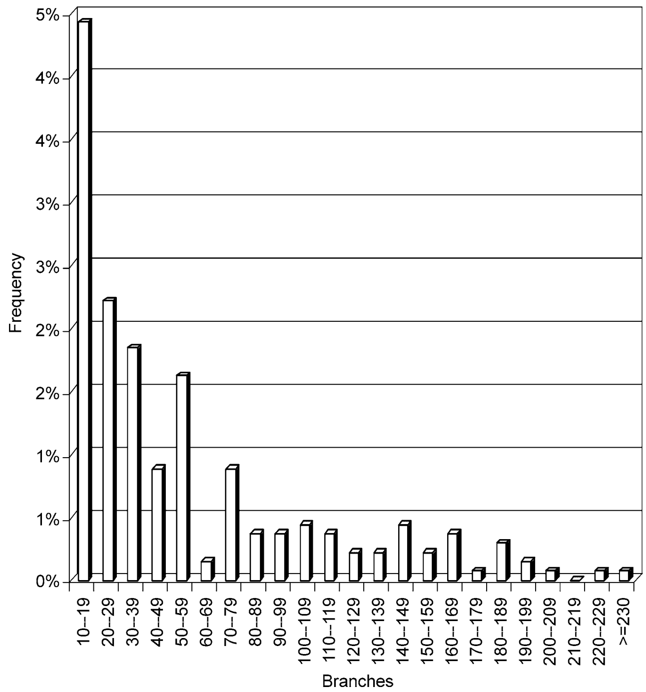


Fig. 2. Sample distribution by number of branches. Observations with 10 or more branches.

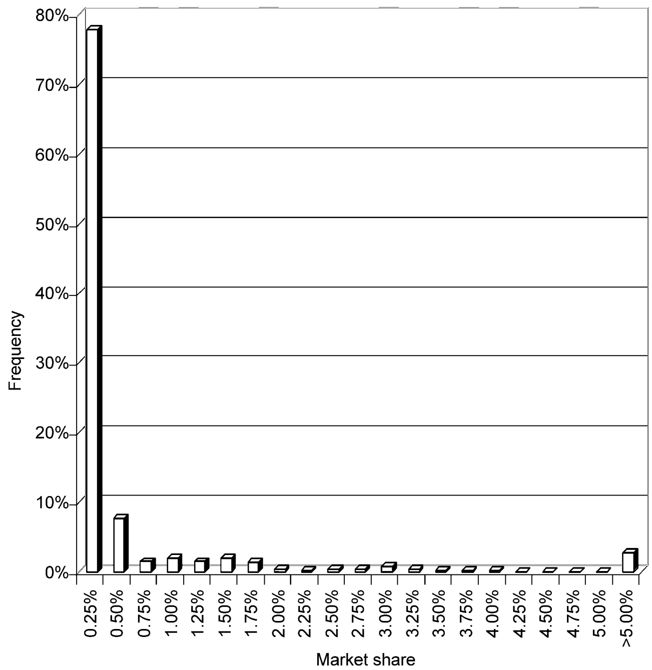


Fig. 3. Sample distribution by market share of loans.

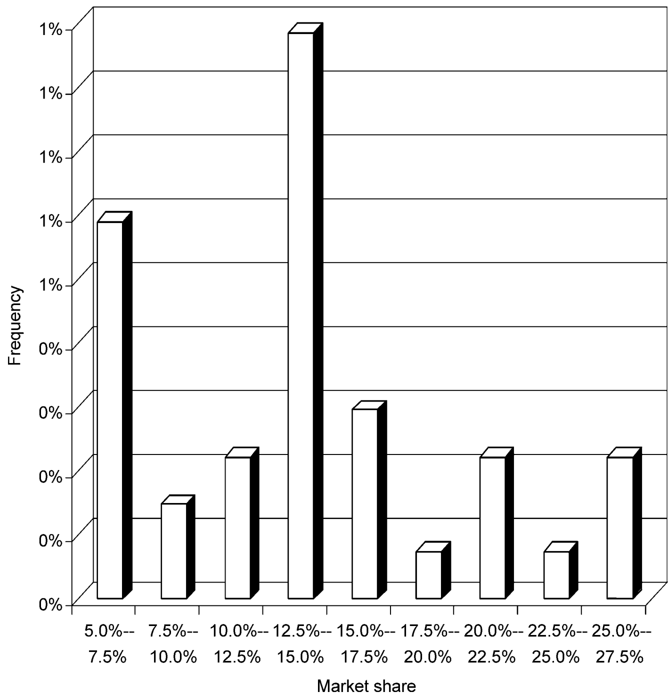


Fig. 4. Sample distribution by market share of loans. Observations with at least 5% market share.

It is evident from these figures that the banking industry is characterized as having many very small banks. About 27% (44%) of the observations have only 1 branch (2 or less branches) and 78% have market share smaller than 0.25%. The largest bank in terms of branches has 240. In terms of market shares, the largest bank captures 27.4%. The distribution of bank size by both measures is fairly stable throughout the period. The correlation between banks' number of branches and market shares is as high as 0.8775. It is noteworthy to mention, however, that even the very large commercial banks with nationwide branch networks supply loans to the typical small bank customer and to small local businesses. Thus, all banks compete in the market for retail lending.

Table 1 shows some more important details of the Norwegian banking industry. The growth in bank lending between 1994 and 1996 is associated with the strong recovery of the Norwegian economy from a recession in the late 1980s and early 1990s. From 1987 to 1993, the average annual growth rate in GDP was just below 2%, whereas the average annual growth rate between 1994 and 1996 was 4.8%. Hence, the increased importance of bank lending at the end of the sample period should be attributed to the relatively strong procyclical pattern of the overall borrowing activity. It follows that to the individual bank this growth may be considered exogenous.

A closer look at the structure of the market for bank loans reveals a relatively stable pattern. The correlations between current market shares and the market shares lagged from one to four years are 0.997, 0.968, 0.935, and 0.927, respectively. These figures may indicate very little switching of borrowers between banks over a one, and even over two year periods. Strictly speaking, high intertemporal correlation among market shares may also emanate from intensive switching among banks resulting in (close to) zero change in net market shares. The challenge of the proposed model is to uncover actual switching hidden in the market shares evolution.

4.2. Construction of variables

In estimating the cost function, loans extended are treated as bank output.[32] Four variable factors of production are specified: labor, materials, physical capital (machinery and buildings) and funding. The latter includes both deposits and borrowed money, i.e., all bank debt except subordinated debt.[33]

Bank- and year-specific prices of labor are computed as total labor costs per man-hour. The price of materials is measured by a price index for material inputs to financial institutions, collected from the national accounts statistics. Thus, it varies through time but constant across banks. That is also the case for the price of physical capital which is proxied by the ten-year interest rate on government bonds.

The price of funding is a weighted average of the interest rate on deposits and the interest rate on borrowed money. The latter is measured by the three-month money market interest rate, thus it only varies across time, not across banks.[34] The interest rate on deposits, however, varies in both dimensions. It is calculated as banks' interest expenditures on deposits divided by the mean of deposits at the beginning and the end of the year.[35] Bank-specific interest rates on loans are measured likewise.

The discount factor, δ_t, applied for estimating the parameters of the first-order condition (3.1) is calculated using the 9 to 12 month interest rate on T-bills, which amounts to an average of 9.16% in the sample period.

As described in the next section, we apply a lead and a lag of three years, leaving us with observations only from 1991, 1992, and 1993, and a total of 411 observations in the sample. In these three years the number of banks varied from 147 to 143.

5. Estimation and results

There are two major issues to be attended in the estimation process. The first relates to the proper length of period in which switching may take place. The second important issue relates to the definition of the geographic and product market, i.e., to account for two aspects:

---

[32] The output aggregation problem in the context of banking is dealt with in Kim (1986).

[33] Subordinated debt is considered quasi equity, partially being counted as part of the capital base when measuring the capital adequacy according the Basel rules. Only a few banks in our sample have issued subordinated debt.

[34] The three-month money market interest rate is represented by the three-month effective NOK eurorate.

[35] Whenever lagged values of physical capital, loans, deposits, or debt are used to calculate interest rates, the lagged values are adjusted for bank mergers, i.e., the bank structure in period t is imposed on the variables lagged to year $t-1$.

(i) customer location preferences and the corresponding banks' branch-network size, and
(ii) the ability of banks to provide the entire continuum of loan sizes demanded.

Regarding the first issue, the loan maturity may help us in deciding on the preferred length of period to be used in the estimation, as we may think of two switching patterns. First, a borrower may terminate the loan agreement and repay the loan to its provider bank whenever sufficiently better loan conditions are offered by a rival bank. Second, a borrower may consider switching to other banks only upon maturity of the loan previously taken. Fixing the model's period equal to the average length of loans will shed light on the latter switching behavior since in this case borrowers have, on average, one chance to switch between lenders in each period.[36] Using this argument, the model's estimation using a shorter period than the average length of loans would involve both switching patterns.

When estimating the model with one- and two-year lags, we find out that borrowers hardly switch. Either the estimated switching costs are of such magnitude that the restriction made to insure the transition probabilities lie in the $[0, 1]$ interval (see (3.8)) was violated (with one-year lag), or the estimated parameters lacked precision (in the two-year lag).

The intertemporal correlation coefficients between market shares reported in Section 4.1 may indicate that there may be more switching of borrowers over a longer time horizon. Given the need for both a lead and a lag of the same length, the longest lag we can use, maintaining the possibility of estimating the whole system simultaneously, is three years.[37] Thus, all our reported results are based on a three-year period.

Dealing with the issue of product–market definition amounts to deciding which banks should be included in the estimation. This issue is dealt with along both dimensions, that of the branch-network size, and that of loan size. If branch-network size (as is conventionally measured by the number of branches) affects the state and degree of competition, due for instance to location preference, then banks which have very few branches may not correspond well to the model's assumption of mutual competition. As there is no clear and natural cutoff number of branches we estimate the equation system on various size subsamples of banks, defined by different *minimal* number of branches and compare the resulting switching cost estimates. This procedure enables us to see whether the various cutoffs result in robust estimates of the switching cost parameters. Table 2 displays the results of the estimation with subsamples using different *minimal* number of branch cutoffs.[38]

As to the issue of geographic market definition, each bank's relevant competitors should be specified. This, however, is too demanding from aggregate data, lacking individuals' switching information. Ignoring that issue may result in an upward bias of the switching

[36] Degryse and van Cayseele (2000) provide evidence on bank relationship within a bank-based system in Belgium. The average time length of loans in their sample is 2.39 years.

[37] Estimating the system with a four-year lag structure produced similar results to the reported three-year lag structure. This, however, was feasible only with a two stage (separate) estimation of the cost function and the rest of the system. A joint estimation of the system with a four-year lag, and estimation using higher-order lags was not feasible due to the short time dimension of the available data.

[38] The switching cost estimates (s) in the table are reported in terms of annual interest rates.

Table 2
Switching cost subsamples by minimal number of branches, parameter estimates

Min. no. of branches:	1[a]	10	20	30	40	50	60
γ_y	0.989	0.919	0.919	0.915	0.899	0.891	0.785
(output)	(0.007)	(0.009)	(0.010)	(0.011)	(0.012)	(0.016)	(0.023)
γ_w	0.139	0.149	0.150	0.148	0.145	0.147	0.145
(labor)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
γ_k	0.018	0.014	0.013	0.013	0.014	0.014	0.014
(capital)	(0.001)	(0.001)	(0.002)	(0.005)	(0.001)	(0.005)	(0.001)
γ_f	0.720	0.708	0.709	0.710	0.714	0.715	0.718
(funding)	(0.003)	(0.004)	(0.005)	(0.005)	(0.016)	(0.005)	(0.005)
α_1	−4.994	−5.043	−5.391	−6.352	−6.711	−6.670	−6.727
(trans. prob. slope)	(0.786)	(0.939)	(1.176)	(0.965)	(0.996)	(1.248)	(0.878)
s	4.120	6.040	6.867	3.361	2.892	4.423	2.142
(switching cost, pct.)	(2.164)	(2.747)	(3.114)	(1.848)	(1.665)	(2.321)	(1.598)
Numbers of obs.	411	61	45	36	30	27	20

The table reports parameter estimates and goodness-of-fit statistics for the jointly estimated equation system consisting of Eqs. (3.1), (3.4), (A.1), and (A.2). Standard errors (in parentheses) are White (1980) heteroschedasticity-adjusted. Second-order terms of the cost function are not reported for brevity of exposition.
[a] Indicates total sample.

cost estimate, the degree of which may be assessed by analyzing the results of the estimation based on different minimal number of branch cutoffs.

As depicted in Table 2, the point estimate of the switching cost estimates range from a low (and insignificant) value of 2.14% for the group consisting of the very large banks (those with 60 or more branches), to a high (and significant) value of 6.87% for the group consisting of banks with 20 or more branches. The point estimate based on the entire sample is 4.12%. We note that all switching cost estimates (except that of the largest group) reported are statistically significant at the 5% level based on a one tailed test.[39] The result that switching cost estimates decrease with bank size may pick up the fact that the larger (branch-network) banks serve a higher proportion of large and mobile (wholesale) customers than smaller retail customers. Large customers are usually publicly-traded firms for which asymmetric information problems are relatively less important, and thus the ‘lemons’ phenomenon is less problematic. Furthermore, larger customers are better at gathering and processing the relevant information in financial markets and therefore are more mobile across banks. This is consistent with Berg and Kim (1998) who document no market power in the market for wholesale loans but strong market power for retail loans.[40]

To verify that the estimates yield valid results, note that the constraint (3.6) is not binding. The point estimate of α_1, the slope of the transition probability functions, ranges from −6.73 for the very large banks to −5.05 for the smaller (less branch intensive) banks.

39 Switching costs cannot be negative.

40 It may also be possible that small borrowers from small banks have longer relationships than those of small borrowers from large banks, indicating that the latter borrowers may have lower switching costs compared to the former. We thank Mitchell Petersen for this insight. Also, Berger et al. (2002) find that larger banks have shorter relationships with their small business borrowers (US data)—possibly consistent with the lower switching costs we find and the lower market power found by Berg and Kim (1998).

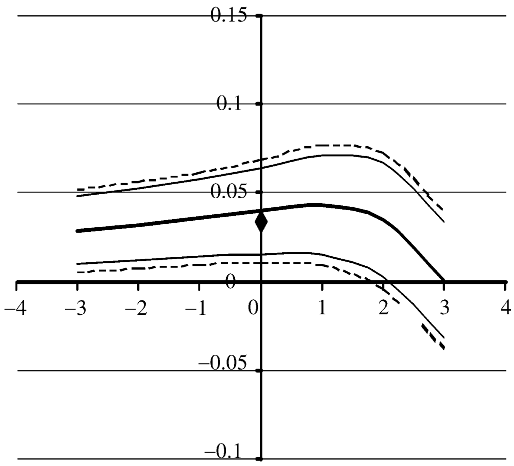


Fig. 5. Switching cost estimates for the subsample of observations with at least 30 branches, with restricted α_1 values. The horizontal axis shows deviations from the unrestricted α_1 point estimate, in standard error units. The diamond plots the unrestricted point estimate of the switching costs. The bold line provides the restricted switching cost estimates. The continuous and dashed lines designate the 90% and the 95% confidence intervals, respectively.

The point estimate for the entire sample is -4.99, all of which are statistically significant at the 1% level. Indeed, (abstract of switching costs) in a market with Bertrand behavior and highly substitutable goods ex-ante, as is the market for banking loans, the price–cost margin should be very low, as implied by this α_1 estimate.[41]

To verify that the switching-cost point estimates are insensitive to the slope of the transition probabilities (α_1), we re-estimate the equation system, imposing various values (of standard error deviations from the point estimate) for the slope of the transition probabilities, thereby yielding conditional switching cost estimates. As an illustration we plot these values as a function of the imposed α_1 values for the case of banks with at least 30 branches (36 observations).[42] This is depicted in Fig. 5.

The horizontal axis of the figure measures deviations from the point estimate of α_1 (which is -6.35) in standard error units. We also provide 90% and 95% confidence intervals (bounded by the solid lines and the dashed lines, respectively) around the point estimates of the conditional s. It is evident from the figure that the point estimate of the switching cost is not very sensitive to changes in α_1. We also reestimate the equation system on each group of observations, as in Table 2, imposing various α_1 values around its point estimate. In Fig. 6 we plot the switching costs estimates for these subsamples by *minimal* number of branches, for the different imposed α_1 values and their standard error deviations. This figure shows that the switching cost estimates are rather insensitive to

[41] As an illustration, the (grand) sample's point estimate of α_1 implies that had switching costs been eliminated (by an exogenous transfer for instance) the average price–cost margin would have been as low as 0.15%. To see that, calculate the price–cost margin in (2.37) for an average size bank.

[42] In what follows we present this exercise for each of the groups in a combined figure (see Fig. 6).

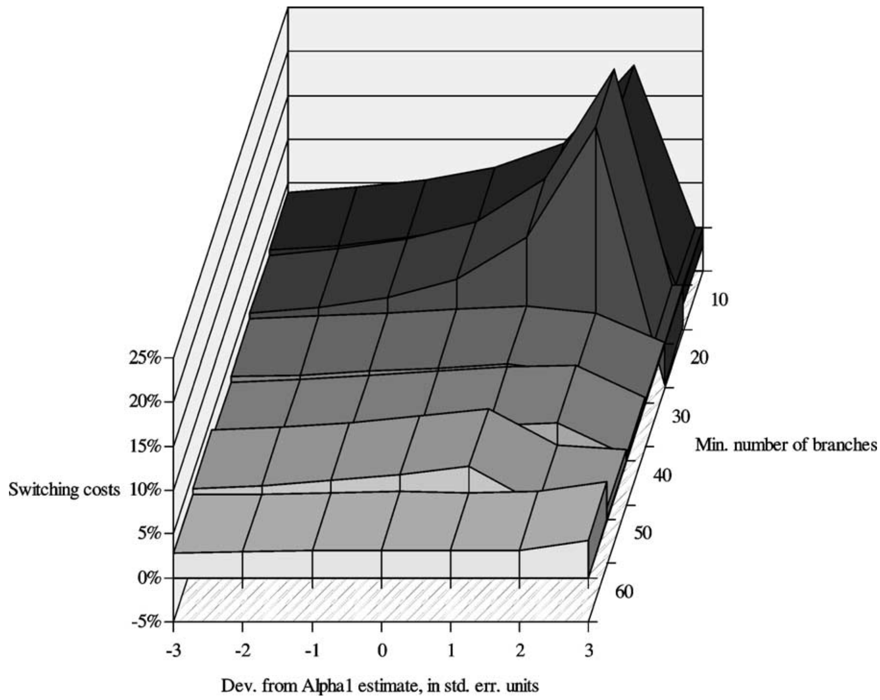


Fig. 6. Switching cost estimates for subsamples by minimum number of branches, with restricted α_1 values.

variation in the values of α_1. The only large deviation occurs when the equation system is estimated on the 10- and 20-branches cutoff samples, and α_1 is restricted to be two standard errors larger than its unrestricted value.

The second important aspect of industry definition is loan size. Specifically, small banks have very low lending capacity resulting in loans of very small sizes whereas large banks have high lending capacity resulting in the provision of the entire continuum of loan quantity demanded. This results in different mix of customers served by the various banks. Thus, relatively large customers (usually firms) will find it impossible to switch to small banks which do not have the capacity to serve them. This may have the effect of introducing very large and perhaps pathological switching costs in the case of very small banks which, in turn may result in quite a large point estimate for the average switching cost when estimation is carried out on the entire sample of banks. Since our data set does not provide information on the customer mix, we have estimated our system of equations on various size subsamples, defined by the *minimal* total loan size.[43] The results are presented in Table 3 and, for expositional purpose also, are depicted in Fig. 7.

The switching cost parameter estimates range from as low, and statistically insignificant, a value of 0.21% for the group consisting of the largest banks (those with loans over 12

[43] The reason for the various cutoff points is the compatibility with those applied to the number of branches in terms of number of observations.

Table 3
Switching cost subsamples by loan size, parameter estimates

Min. loans (bill.)	> 0[a]	2	3	4	5.5	7.5	12
γ_y	0.989	0.946	0.944	0.942	0.930	0.925	0.748
(output)	(0.007)	(0.013)	(0.012)	(0.013)	(0.012)	(0.016)	(0.020)
γ_w	0.139	0.148	0.149	0.149	0.147	0.146	0.146
(labor)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
γ_k	0.018	0.014	0.014	0.014	0.014	0.015	0.015
(capital)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
γ_f	0.720	0.715	0.714	0.713	0.713	0.715	0.717
(funding)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
α_1	−4.994	−4.699	−5.551	−5.838	−6.748	−6.725	−6.679
(trans. prob. slope)	(0.786)	(1.233)	(1.235)	(1.341)	(1.166)	(1.490)	(0.728)
s	4.120	8.437	6.341	6.270	4.121	5.690	0.213
(switching cost, pct.)	(2.164)	(4.363)	(3.156)	(3.279)	(2.089)	(2.948)	(1.309)
Number of obs.	411	55	48	42	36	28	21

The table reports parameter estimates and goodness-of-fit statistics for the jointly estimated equation system consisting of Eqs. (3.1), (3.4), (A.1), and (A.2). Standard errors (in parentheses) are White (1980) heteroschedasticity-adjusted. Second-order terms of the cost function are not reported for brevity of exposition.

[a] Indicates total sample.

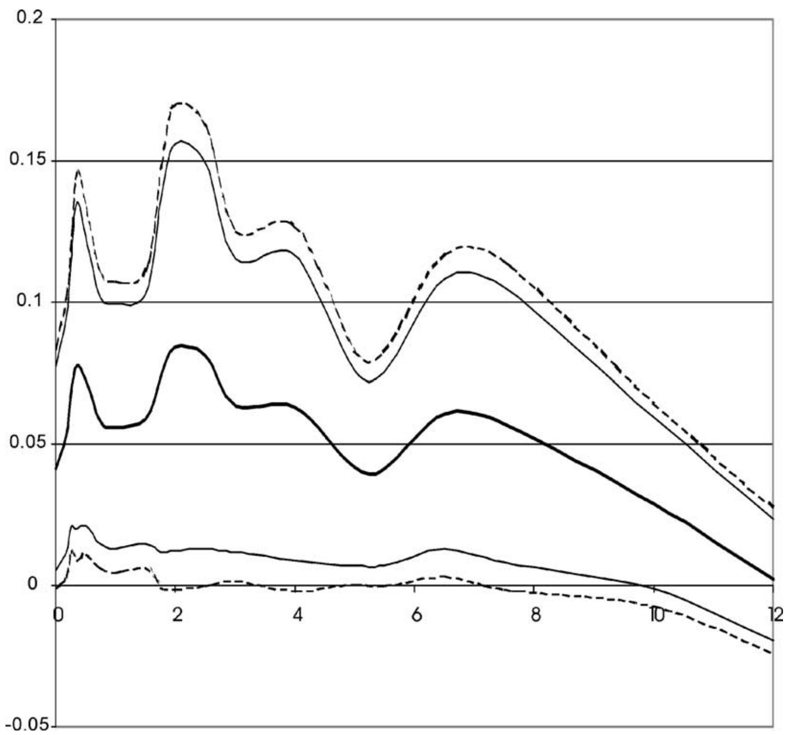


Fig. 7. Switching cost estimates for subsamples by loan size market share. The bold line provides the switching costs estimates. The continuous and dashed lines designate the 90% and the 95% confidence intervals, respectively.

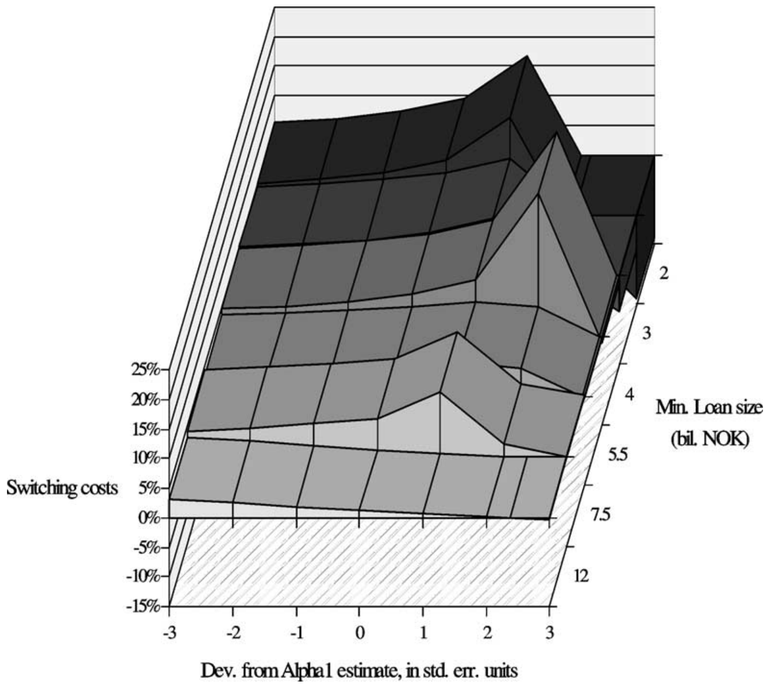


Fig. 8. Switching cost estimates for subsamples by minimum loan size, with restricted α_1 values.

bill. NOK) to, as high and statistically significant, a value of 8.44% for the group consisting also of the smaller banks. These results should be juxtaposed on the point estimate based on the entire sample (4.12%, which was discussed earlier).

The general pattern here is very similar to the one present when estimation was based on the size of the branch network. This should not come as a surprise as the correlation between size of banks measured by the branch-network size and size of loans is quite high (0.8775 for the entire sample). As we did for the estimation on subsamples by minimal number of branches, we check the sensitivity of the switching cost estimates to the slope of the transition probability. In Fig. 8 we plot the switching cost estimates for subsamples by loan size, for different imposed α_1 values and their standard error deviations.

As before, the various switching cost estimates are insensitive to variations in α_1. Large deviations occur only when the equation system is estimated on observations with loans of at least 2 bil. NOK when α_1 is restricted to be one standard error larger than its unrestricted value and on observations with loans of at least 4 bil. NOK when α_1 is restricted to be two standard errors larger than its unrestricted value.

We reiterate that the estimated switching costs described above capture not only the direct pecuniary effect but also the various psychological aspects of switching, as well as the indirect, and perhaps the most important, economic costs associated with the loss of the capitalized value of long-term customer bank relationship. One major aspect contributing to the relatively high switching cost estimates is the one related to the downturn of the economy resulting in system wide banking problems which exacerbated the 'lemons' problem and hence the reluctance of banks to provide loans to new customers.[44] The other major aspect has to do with the long duration of bank–customer relationship documented in American and European banking.

5.1. Counterfactuals

The parameter estimates reported in the previous section can be used to extract a host of interesting quantitative information regarding the effects of switching costs on both borrowers' and banks' characteristics. In what follows we present and discuss characteristics such as:

(i) the average duration of contact between a bank and its borrowers,
(ii) the probability of staying with a particular bank vs. that of switching,
(iii) the contribution of previous periods' market shares to the current ones, and
(iv) the contribution of the lock-in phenomenon to banks' marginal profits.

Inserting the (total sample) estimates of α_1 and s into (3.5) we find that the probability that an existing customer will continue to purchase from the firm is larger than the probability that a rival's customer will switch to the firm, by 20% (see Table 4 for the probabilities of staying and switching). A closer look at this probability when markets are defined according to the size of the branch network generates a range of 17–32%, which averages 30%. Defining the market according to loan size, this range is 2–43% with an average of 35% (see sixth column in Table 4).[45]

For the average bank, we may assess the transition probabilities[46] (2.12) and (2.13):

$$\underset{i\to i,t}{\widehat{\Pr}} = \bar{\alpha}_0^i + \hat{\alpha}_1\big(p_{i,t} - \bar{p}_{iR,t} - \hat{s}\big) \tag{5.1}$$

and

$$\underset{iR\to i,t}{\widehat{\Pr}} = \bar{\alpha}_0^i + \hat{\alpha}_1\left(p_{i,t} - \bar{p}_{iR,t} + \frac{\hat{s}}{n-1}\right), \tag{5.2}$$

where $\bar{\alpha}_0^i = 1/n$. The estimated probabilities in (5.1) and (5.2) indicate that after one year 40% of the average bank's customers switch to other banks. Recall that the transition probabilities are based on a three-year period. We annualize these probabilities as follows: $\theta = 1 - (\widehat{\Pr}_{i\to i,t})^{1/3} = 0.40$ where θ is the annual defection rate (last column in Table 4). Of course, these estimates are based on the total sample.[47] Using the various market definitions, as above, we get similar results but realize that for the larger banks the

[44] It should be noted, however, that during the problematic period (1990–1993) only one small bank was actually closed. All other banks that got into problems were able to continue their lending operations due to government interventions.

[45] Note that Eq. (3.5) = Eq. (2.15).

[46] Hats over variables ($\hat{}$) denote estimates.

[47] While the intertemporal correlations of market shares are measures of net customer flow (arrivals minus departures), the model estimates reveal the actual switching. For instance, there may exist intensive switching even when the intertemporal correlation equals unity.

Table 4
Characteristics by subsamples

Characteristic		No. obs.	$\widehat{\Pr}_{i\to i,t}$	$\widehat{\Pr}_{iR\to i,t}$	$\lambda_{k=0.99}$	$\frac{\partial\sigma_{i,t}}{\partial\sigma_{i,t-1}}$	$\frac{mvl_{i,t}}{\partial V_{i,t}/\partial y_{i,t}}$	θ
By branch	10	61	0.35	0.03	13.3	0.32	0.25	0.29
	20	45	0.44	0.04	16.7	0.40	0.30	0.24
	30	36	0.30	0.06	11.4	0.24	0.18	0.33
	40	30	0.29	0.08	11.3	0.21	0.17	0.33
	50	27	0.41	0.07	15.3	0.34	0.26	0.26
	60	20	0.29	0.12	11.3	0.17	0.13	0.34
	Average		0.35	0.05	13.6	0.30	0.23	0.30
By loan size	2	55	0.45	0.03	17.3	0.42	0.32	0.23
	3	48	0.41	0.04	15.7	0.37	0.29	0.25
	4	42	0.44	0.04	16.7	0.40	0.30	0.24
	5.5	36	0.36	0.06	13.6	0.30	0.23	0.29
	7.5	28	0.49	0.06	19.4	0.43	0.33	0.21
	12	21	0.16	0.14	7.50	0.02	0.01	0.46
	Average		0.38	0.03	13.48	0.35	0.23	0.28
Total sample		411	0.21	0.01	8.90	0.20	0.16	0.40

$\lambda_{k=0.99}$ represents the duration of relationship based on the time required for k% of the customers to switch. The term $\partial\sigma_{i,t}/\partial\sigma_{i,t-1}$ represents the contribution of the $t-1$ market share to the market share in period t. The term $mvl_{i,t}/(\partial V_{i,t}/\partial y_{i,t})$ represents the proportion of the marginal value of locked-in customer in the marginal increase in the bank's present value. The θ is annual defection rate.

$\widehat{\Pr}_{i\to i,t}$ are of smaller magnitude. This is consistent with the fact that larger banks have higher proportion of wholesale customers over which market power is indeed minimal, see Berg and Kim (1998). The turnover period (the average duration of customer–bank relationship) is thus, $3 \div (1-0.213) = 3.812$ years. Turnover calculations, however, are somewhat misleading since, in fact, some of the bank's customers may never switch. Thus, in addition to the turnover period, we may calculate the time required for k percent of the customers to switch. Denoting this time period by $\lambda = \ln(1-k) \div \ln(1-0.40)$, for $k = 99\%$, $\lambda = 8.9$ years. Again, these calculations are based on the estimation using the entire sample observations. However, as was demonstrated, results are somewhat different when estimation is carried out on different subsamples according to branch-network size or output size. The range of customer–bank relationship is between 11.3 and 16.7 years with an average of 13.6 years if markets are defined by branch-network size, or between 7.5 and 19.4 years with an average of 13.5 years if the definition is according to loan size. The interesting point though is the general pattern that the duration of relationship is smaller for the large banks (7.5 years for those with loans over 12 billion NOK, see Table 4), again consistent with Berg and Kim (1998).

Our finding is very much in line with recent literature on relationship lending within a bank-based system using detailed (survey) micro data. Ongena and Smith (2001) use panel data of connections between Oslo Stock Exchange-listed firms and their banks, for the 1979–1995 period. After correcting for censoring bias, the estimated mean duration of firm-bank relationship varies between 15 and 18 years. This is also in line with other studies using European and American survey data. Angelini et al. (1997) report an average duration of 14 years for Italy. Cole (1998) finds mean duration of US firms (based on data from the National Survey of Small Business Finance) to be 7 years. Harhoff and Körting (1997) report average duration of 13 years for German firms. Petersen and Rajan (1994) report an 11 years mean duration in the USA. Degryse and van Cayseele (2000) report a mean value for the length of bank relationship in Belgium of 7.8 years.

The estimated parameters also suggest (sixth column of Table 4) that the contribution of the previous period's market share to the current one, as defined by (2.16) is 0.2 when considering the entire sample. This means that about 20% of the average bank's market share is due to its bank–borrower relationship in the previous period. This however, can be as low as 17% and 0.2% for the very large banks when the market is defined by the branch-network size and by the loan size, respectively. When smaller banks are also included it is 32% and 42%, respectively. Again, this points to the higher mobility of (larger) customers working with larger banks where output is characterized by a higher proportion of wholesale loans.

An interesting and important issue is that of the proportion of the marginal value of a locked-in customer $(mvl_{i,t})$ to the marginal increase of the bank's present value which is due to an additional locked-in customer $(\partial V_{i,t}/\partial y_{i,t})$, as in (2.39). Based on the total sample, the value of $mvl_{i,t}/(\partial V_{i,t}/\partial y_{i,t})$ is 0.16, that is 16% of the customer's added value is attributed to the lock-in phenomenon generated by switching costs. We note, however, that a similar pattern emerges here as it did with the other characteristics considered, namely, the contribution of locked-in customers to banks' value decrease as the size of bank increases. Specifically, as is apparent from Table 4, the contribution of locked-in customers to banks' value ranges from a low of 13% for the very large banks to 30% for the group including also the smaller ones when the market is defined according to branch-network size. It is more pronounced when considering loan size as the definition of the market, where locked-in customers contribute only 1% in the case of the large banks, and 32% in the case of the smaller banks. Once more, these results emanate from the higher proportion of mobile wholesale customers of larger banks. It is noted that on the average (for either definition of the market) the locked-in customers contribute 23% to banks' value.

6. Concluding remarks

We have proposed an empirical model of firms' strategic behavior in the presence of switching costs. Customers' transition probabilities embedded in firms strategic interaction were used in a multiperiod model to derive estimable equations of a first-order condition and market share (demand) equations. The novelty of the proposed model is in its ability to extract information on both the magnitude and significance of switching costs, as well as on customers' transition probabilities, from conventionally-available, highly-aggregated data which do not contain customer-specific information. As a matter of illustration, the model was applied to panel data of banks, to estimate the switching costs in the market for bank loans. We have found that the grand average point estimate of switching costs is about 4.1%, and may be as low as 0.2% when only banks with the largest loan portfolio are included in the definition of the market. When the market is defined according to the branch-network size the switching cost among the largest banks is about 2.1%. On average, 23.0% of the customer's added value is attributed to the lock-in phenomenon generated by switching costs. As much as 35.0% of the average bank's market share is due to its established bank–borrower relationship. The model estimates imply an average duration of bank–customer relationship of 13.5 years. All the above characteristics exhibit lower values for the group of larger banks whose loan portfolio is dominated by more mobile wholesale customers.

To summarize, switching costs in the market for bank loans are quite substantial and constitute a significant portion of the value of a marginal customer to the average firm. The presented technique may be applied to other markets in order to gain insight into the empirical regularity of switching costs.

Acknowledgments

We are grateful to Emanuel Barnea, Pedro Barros, Benjamin Bental, Luís Cabral, Clive Granger, Giora Hanoch, Patricia Jackson, Eugene Kandel, Benny Levikson, Loretta Mester, Espen Moen, Dan Peled, Christian Riis, Asbjørn Rødseth, Agnar Sandmo, Steven Sharpe, Lars Sørgard, Erling Steigum, Siri Pettersen Strandenes, Dag Tjøstheim, Ana Lozano-Vivas and Yitzhak Weit for their comments. We thank participants of the Annual Research Meeting of the Norwegian Economic Association and workshop participants at Ben Gurion University, Humboldt University, Center for Financial Studies Goethe University, Norges Bank, Tel Aviv University, Central Bank of Israel, Hebrew University, Norwegian School of Management, University of Oslo, Tinbergen Institute, University of Malaga, Norwegian School of Economics, UMH Workshop on Banking and Finance Allicante Spain, the CEPR conference in Lisbon, DIW-Berlin, WZB-Berlin, the Conference on Consumer Transaction and Credit at the Federal Reserve Bank of Philadelphia and Basel Committee workshop on empirical banking research. Finally, we are indebted to Mitchell Petersen (editor) and an anonymous referee for very extensive review and comments which have much improved the final result.

Appendix A. The case of noisy prices

Denote the firm's true price by $p^*_{i,t}$, and rivals' true price average by $\bar{p}^*_{iR,t}$. The probability that firm i's customer will continue to purchase from it then becomes:

$$\Pr_{i\to i,t} = A^i_0 + \alpha_1 p^*_{i,t} + \alpha_2\big(\bar{p}^*_{iR,t} + s\big). \tag{A.1}$$

The firm's and the rivals' *observed* prices equal the true prices plus error terms:[48]

$$p_{i,t} = p^*_{i,t} + \xi_i \tag{A.2}$$

and

$$p_{j,t} = p^*_{j,t} + \xi_j. \tag{A.3}$$

[48] Notice that the error terms may have any specific mean or variance.

When inserting the true prices into the transition probability function all error terms are captured by the intercept, and no effect on α_1, α_2 and s remains:

$$\begin{aligned}\Pr_{i\to i,t} &= A_0^i + \alpha_1 p_{i,t}^* + \alpha_2\big(\bar{p}_{iR,t}^* + s\big)\\ &= A_0^i + \alpha_1(p_{i,t} + \xi_i) + \alpha_2\big(\bar{p}_{iR,t} + \bar{\xi}_{iR} + s\big),\end{aligned} \tag{A.4}$$

where $\bar{p}_{iR,t}^*$ and $\bar{\xi}_{iR}$ are the average of rivals' prices and error terms, respectively. The transition probability then becomes (2.8), where $\alpha_0^i = A_0^i + \alpha_1\xi_i + \alpha_2\bar{\xi}_{iR}$. The same applies to $\Pr_{iR\to i,t}$ in (2.9).

Appendix B. The cost function

Bank i's time-variant cost function is represented by its second-order Taylor-series approximation of the following form:

$$\begin{aligned}\ln c_{i,t} &= \gamma_0 + \gamma_y \ln y_{i,t} + \sum_k \gamma_k \ln w_{k,i,t}\\ &\quad + \frac{1}{2}\left(\sum_k\sum_l \gamma_{kl} \ln w_{k,i,t} \ln w_{l,i,t} + \gamma_{yy}(\ln y_{i,t})^2\right)\\ &\quad + \sum_k \gamma_{ky} \ln w_{k,i,t} \ln y_{i,t}.\end{aligned} \tag{B.1}$$

Application of Shepherd's lemma to (B.1) yields the following factor demand (share) equations:

$$\frac{\partial \ln c_{i,t}}{\partial \ln w_{k,i,t}} \equiv m_{k,i,t} = \gamma_k + \sum_l \gamma_{k,l} \ln w_{l,i,t} + \gamma_{ky} \ln y_{i,t}, \tag{B.2}$$

where $m_{k,i,t}$ is the share of the kth factor in bank i's period-t production cost. The following restrictions for symmetry and linear homogeneity in prices are imposed on the cost (production) structure:

$$\gamma_{kl} = \gamma_{lk}, \quad \forall k, l; \qquad \sum_k \gamma_k = 1; \qquad \sum_k \gamma_{ky} = 0; \qquad \sum_l \gamma_{kl} = 0, \quad \forall k. \tag{B.3}$$

Marginal production cost estimates are obtained from (B.1):[49]

$$mc_{i,t} \equiv \frac{\partial c_{i,t}}{\partial y_{i,t}} = \frac{c_{i,t}}{y_{i,t}}\left(\gamma_y + \gamma_{yy} \ln y_{i,t} + \sum_k \gamma_{ky} \ln w_{k,i,t}\right) \tag{B.4}$$

and is inserted in the first-order condition (3.1) in the text.

[49] Due to Shephard's Lemma, one of the input share equations is omitted to insure non-singularity.

Appendix C. Differentiating the firm's output

The demand faced by the firm, after substituting the linear transition probabilities is

$$y_{i,t} = \Big(y_{i,t-1}\big(\alpha_0^i + \alpha_1(p_{i,t} - \bar{p}_{iR,t} - s)\big) + y_{iR,t-1}\Big(\alpha_0^i + \alpha_1\Big(p_{i,t} - \bar{p}_{iR,t} + \frac{s}{n-1}\Big)\Big)\Big)g_t \tag{C.1}$$

or

$$y_{i,t} = \Big(-y_{i,t-1}\frac{n}{n-1}s\alpha_1 + y_{t-1}\alpha_0^i + y_{t-1}\alpha_1\Big(p_{i,t} - \bar{p}_{iR,t} + \frac{s}{n-1}\Big)\Big)g_t, \tag{C.2}$$

where $y_{t-1} = \sum y_{i,t-1}$.

Differentiating w.r.t. the firm's price and rivals' average price gives

$$\frac{\partial y_{i,t}}{\partial p_{i,t}} = y_{t-1}\alpha_1 g_t \tag{C.3}$$

and

$$\frac{\partial y_{i,t}}{\partial \bar{p}_{iR,t}} = -y_{t-1}\alpha_1 g_t. \tag{C.4}$$

The time $t+1$ demand is

$$y_{i,t+1} = \Big(-y_{i,t}\frac{n}{n-1}s\alpha_1 + y_t\alpha_0^i + y_t\alpha_1\Big(p_{i,t+1} - \bar{p}_{iR,t+1} + \frac{s}{n-1}\Big)\Big)g_{t+1}. \tag{C.5}$$

Differentiating w.r.t. the firm's time t price and rivals' time $t+1$ average price gives

$$\frac{\partial y_{i,t+1}}{\partial p_{i,t}} = -\frac{\partial y_{i,t}}{\partial p_{i,t}}\frac{n}{n-1}s\alpha_1 g_{t+1} = -y_{t-1}\alpha_1\frac{n}{n-1}s\alpha_1 g_t g_{t+1} \tag{C.6}$$

and

$$\frac{\partial y_{i,t+1}}{\partial p_{i,t+1}} = y_{t-1}\alpha_1 g_t g_{t+1}. \tag{C.7}$$

References

Angelini, P., Di Salvo, R., Ferri, G., 1997. Availability and cost of credit for small businesses: Customer relationships and credit cooperatives. Mimeo, Banca d'ltalia.

Ausubel, L., 1991. The failure of competition in the credit card market. The American Economic Review 81, 50–81.

Beggs, A., Klemperer, P., 1992. Multi-period competition with switching costs. Econometrica 60, 651–666.

Berg, S.A., Kim, M., 1998. Banks as multioutput oligopolies: An empirical evaluation of the retail and corporate banking markets. Journal of Money, Credit, and Banking 30, 135–153.

Berger, A., Miller, N., Petersen, M., Rajan, R., Stein, J., 2002. Does function follow organizational form? Evidence from the lending practices of large and small banks. Kellog working paper, Northwestern University, Evanston, IL.

Boot, A.W., 2000. Relationship banking: What do we know? Journal of Financial Intermediation 9, 7–25.

Borenstein, S., 1991. Selling costs and switching costs: explaining retail gasoline margins. RAND Journal of Economics 22, 354–369.

Cabral, L., Greenstein, S., 1990. Switching costs and bidding parity in government procurement of computer system. Journal of Law, Economics and Organization 6, 453–469.

Caminal, R., Matutes, R., 1990. Endogenous switching costs in a duopoly model. International Journal of Industrial Organization 8, 353–373.

Chen, Y., Rosenthal, R., 1996. Dynamic duopoly with slowly changing customer loyalties. International Journal of Industrial Organization 14, 269–296.

Cole, R., 1998. The importance of relationships to the availability of credit. Journal of Banking and Finance 22, 959–977.

Dahlby, B., West, D., 1986. Price dispersion in an automobile insurance market. Journal of Political Economy 94, 418–438.

Degryse, H., van Cayseele, P., 2000. Relationship lending within a bank based system: evidence from European small business data. Journal of Financial Intermediation 9, 90–109.

Greenstein, S., 1993. Did installed base give an incumbent any (measurable advantage in federal computer procurement. The RAND Journal of Economics 24, 19–39.

Harhoff, D., Körting, T., 1997. Lending relationships in Germany: Empirical results from survey data. Mimeo, Wissenschaftszentrum Berlin.

Hughes, J.P., Mester, L.J., 1998. Bank Capitalization and cost: evidence of scale economies in risk management and signaling. The Review of Economics and Statistics 80, 314–325.

Israel, M., 2001. Tenure dependence in departures from consumer–firm relationships: The case of automobile insurance. Working Paper 0019, The Center for The Study of Industrial Organization, Northwestern University.

James, C., 1987. Some evidence on the uniqueness of bank loans. Journal of Financial Economics 19, 217–235.

Kim, M., 1986. Banking technology and the existence of a consistent output aggregate. Journal of Monetary Economics 18, 181–195.

Klemperer, P., 1985. Markets with consumer switching costs. Quarterly Journal of Economics 102, 375–394.

Klemperer, P., 1987. Entry deterrence in markets with consumer switching costs. Economic Journal 97, 99–117.

Klemperer, P., 1995. Competition when consumers have switching costs: An overview with applications to industrial organization macroeconomics, and international trade. Review of Economic Studies 62, 515–539.

Knittel, C.R., 1997. Interstate long distance rate: Search costs, switching costs, and market power. Review of Industrial Organization 12, 519–536.

Ongena, S., Smith, D.C., 2000. Bank relationships: a review. In: Zenioz, S.A., Harker, P. (Eds.), The Performance of Financial Institutions. Cambridge University Press, Cambridge, UK.

Ongena, S., Smith, D.C., 2001. The duration of bank relationships. Journal of Financial Economics 61, 449–475.

Padilla, A.J., 1992. Mixed pricing in oligopoly with consumer switching costs. International Journal of Industrial Organization 10, 393–411.

Padilla, A.J., 1995. Revisiting dynamic duopoly with consumer switching costs. Journal of Economic Theory 67, 520–530.

Petersen, M., Rajan, R., 1994. The benefits of lending relationships: evidence from small business data. Journal of Finance 49, 3–37.

Rajan, R., 1992. Insiders and outsider: The choice between informed and arm's-length debt. The Journal of Finance 47, 1367–1400.

Schlesinger, H., von der Schulenburg, J.M.G., 1993. Consumer information and decisions to switch insurers. The Journal of Risk and Insurance 60, 591–615.

Selten, R., 1965. Spieltheoretische behandlung eines oligopolmodelles mit nach-frageträgheit. Zeitschrift für die gesamte Staatswissenschaft 121, 301–324.

Shapiro, C., Varian, H., 1998. Information Rules: A Strategic Guide to Network Economy. Harvard Business School Press, Cambridge, MA.

Sharpe, S.A., 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. Journal of Finance 45, 1069–1087.

Sharpe, S.A., 1997. The effect of consumer switching costs on prices: A theory and its application to the bank deposit market. Review of Industrial Organization 12, 79–94.

Shum, M., 1999. Advertising and switching behavior in the breakfast cereals market. Unpublished manuscript, University of Toronto.

Slade, M.E., 1991. Strategic pricing with customer rationing: The case of primary metals. Canadian Journal of Economics 24, 70–100.

Stango, V., 1998. Price dispersion and switching costs: Evidence from the credit card market. Mimeo, Department of Economics, University of Tennessee.

Vale, B., 1993. The dual role of demand deposits under asymmetric information. Scandinavian Journal of Economics 95, 77–95.

von Thadden, E.L., 1998. Asymmetric information, bank lending, and implicit contracts: the winner's curse. Mimeo, DEEP, University of Lausanne.

von Weizsäcker, C., 1984. The cost of substitution. Econometrica 52, 1085–1116.

ELSEVIER

Journal of Financial Economics 61 (2001) 449–475

JOURNAL OF Financial ECONOMICS

www.elsevier.com/locate/econbase

The duration of bank relationships☆

Steven Ongena[a], David C. Smith[b,*]

[a] *Department of Finance, Tilburg University and CentER, Tilburg, NL 5000 LE, Netherlands*
[b] *Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC 20551, USA*

Received 2 April 1998; accepted 27 June 2000

Abstract

We analyze the duration of bank relationships using a unique panel data set of listed firms and their banks from the bank-dominated Norwegian market. We find that firms are more likely to leave a bank as the relationship matures. Small, profitable, and highly leveraged firms maintain shorter bank relationships, as do firms with multiple bank relationships. These findings are robust to censoring, alternate specifications for the distribution of relationship duration, and other control variables relevant to the

☆The views of this paper are the authors' and do not reflect those of the Board of Governors or its staff. For comments, we thank Mitch Berlin, Erik Berglöf, Øyvind Bøhren, Yehning Chen, Dan Deli, Bjørn Eraker, Thomas Hartmann-Wendels, Joel Houston, Jason Karceski, Beni Lauterbach, Melanie Patrick, Jeffrey Pontiff, Jay Ritter, Kristian Rydqvist, Asbjørn Rødseth, Dirk Schiereck, William Schwert (editor), Steven Sharpe, Per Strömberg, Adrian Tschoegl, and Bent Vale. An anonymous referee and Mark Flannery were especially helpful in improving the paper. We also thank seminar participants at the Wharton Conference on the Performance of Financial Institutions (Philadelphia), the NYU-CFS Conference on Credit Risk Management and Relationship Lending (Frankfurt), the 1998 NTU International Conference on Finance (Taipei), the 1998 American (Chicago), 1997 European (Vienna), and 1997 German (Mannheim) Finance Association Meetings, the 1997 Nordic Finance Conference (Stockholm), FIBE XV (Bergen), the Federal Reserve Bank of Cleveland, Federal Reserve Board, Norwegian School of Management, Norwegian School of Economics and Business Administration, Michigan State University, Texas A&M, and Universities of Binghamton, Delaware, New Hampshire, Oslo, and Washington. Qinglei Dai, Didrik von Haffenbrädl, Morten Josefsen, Therese Jørgensen and Oliver Olsen provided research assistance. We received partial support from the Fund for the Advancement of Bank Education and the Research Program in Competitive Enterprises at the Norwegian School of Management.

*Corresponding author.

E-mail address: david.c.smith@frb.gov (D.C. Smith).

PII: S0304-405X(01)00069-1

Norwegian market. Overall, our results cast doubt on theories suggesting that firms become locked into bank relationships.

JEL classification: G21; C41

Keywords: bank relationships; hazard models; duration analysis

1. Introduction

Researchers argue that private information occupies a central place in bank–customer relationships, and that the extent of private information increases as a bank deals with any given customer over time. From the firm's perspective, strong bank relationships are generally considered valuable assets because they lower the cost and increase the availability of credit. For example, banks will be more willing to make unprofitable loans to customers during difficult financial times when they trust losses will be recouped over the course of a long relationship. Firms also gain when a good payment record guarantees future improvements in the conditions of their loan contracts, or as they acquire a valuable reputation as trustworthy borrowers. Diamond (1991), Rajan (1992), Berglöf and von Thadden (1994), Boot and Thakor (1994), Chemmanur and Fulghieri (1994), and von Thadden (1995) formally model the benefits of bilateral and long-term relationships. These aspects of bank–firm relationships are particularly important in the bank-centered financial systems commonly found in continental Europe (including Scandinavia). At the same time, however, the bank's acquisition of private information over the course of a relationship could effectively "lock-in" customers and permit their banks to extract monopoly rents. Greenbaum et al. (1989), Sharpe (1990), Rajan (1992), Petersen and Rajan (1995), and von Thadden (1998) analyze the value of bank relationships in the presence of such holdup costs.

The duration of bank–borrower relationships should reflect the net impact of these offsetting factors. If the value of a relationship tends to increase through time, or if firms become locked into specific banks, then the likelihood of terminating a relationship should decline through time. On the other hand, if relationships generally become less valuable over time and if switching costs are not prohibitively large, we should observe that terminations increase with time. Moreover, it seems likely that a relationship's value will depend on a firm's specific characteristics. Firms facing large information asymmetries with outside investors stand to benefit most from long-term bank relationships but are also particularly susceptible to holdup problems and high switching costs.

Event studies clearly indicate that bank relationships have value. In contrast to other kinds of external financing, bank loan announcements generate positive abnormal returns to the borrower's stock. Mikkelson and Partch (1986), James (1987), Lummer and McConnell (1989), James and Wier (1990), Best and Zhang (1993), Preece and Mullineaux (1994), Billett et al. (1995), and Shockley and Thakor (1998) provide examples of these positive abnormal returns. Slovin et al. (1993) establish the value of bank–borrower relationships in another context, by showing that borrowers earned negative (positive) abnormal returns as Continental Illionois Bank's probability of failure rose (fell). Hoshi et al. (1991) provide indirect evidence about a bank relationship's value. They find that the investment patterns of bank-affiliated firms are less correlated with their cash flows, presumably because intra-*keiretsu* lending is more flexible than the arm's-length financing available to unaffiliated firms.

While these studies show that bank relationships can increase borrower value, they provide little insight into the sources of value for different types of firms. Petersen and Rajan (1994) and Berger and Udell (1995) demonstrate explicitly that firms in longer relationships receive more bank credit and pay lower interest rates on loan commitments. However, their data source (the 1988–89 National Survey of Small Business Finance) provides only limited information about the duration of relationships, and no information about when firms terminate relationships, and under what circumstances.

We utilize a unique data source to study the value of long-term bank relationships, emphasizing the willingness and ability of firms to terminate relationships. Listing regulations require firms on the Oslo Stock Exchange (OSE) to report annually the banks that they deal with on a regular basis. We track these relationships from 1979 to 1995, recording when firms start and end relationships, and when they switch from one bank to another. Norway provides a natural setting for studying the value of relationships because banks supply 90% of all commercial debt.[1] We combine each relationship history with financial information about the firm and its banks, then use hazard function estimators to infer the determinants of relationship duration. The analysis is complicated by the fact that the data source does not include the complete history of every relationship. Relationships started prior to 1979 or before a firm lists on the OSE are left censored, while relationships that continue beyond 1995 or after a firm delists are right censored. We analyze the sensitivity of our results to corrections for left and right censoring. Because of these corrections, our paper also provides an example of the value of censored-robust estimators.

[1] Commercial debt consists of loans to all non-financial firms by financial institutions. Source: *Statistical Yearbook of Norway, 1996*. Although bank-dominated on the debt side, regulations in Norway forbid a bank from taking significant equity positions in non-financial companies. As of year-end 1994, Norwegian banks owned less than 1% of the equity in the non-financial sector Nilsen, (1995).

We find that the probability of ending a bank relationship increases in duration, suggesting that the value of relationships decline over time. Moreover, small, young, and highly leveraged firms maintain the shortest relationships. In other words, those borrowers commonly thought to be most dependent on bank financing terminate their relationships quickly. These firms do not appear to be forced out by their banks, nor do they leave because of lending capacity constraints, mergers, or bank distress. Taken together, our results suggest that firms do not become locked into bank relationships. Moreover, we show that bank competition has two opposing effects on relationship duration. Multiple-bank firms terminate relationships earlier than single-bank firms. But the multiple-bank firms tend to turn over newer relationships and keep one long-term relationship. Therefore, long-term relationships appear valuable to firms that are unlikely to face credible holdup threats from one monopolistic bank.[2]

Despite the data's richness, we do not have direct information on why relationships begin or end, nor can we observe information specific to each relationship, such as the credit conditions of the loan contract or the non-lending services provided by the bank. Instead, we must base our inferences on what is revealed through the observed terminations. This limits our overall conclusions. For instance, we cannot measure the gains or losses to customer welfare over the course of the relationships. Nevertheless, we believe that our data set, combined with our methodology, provides interesting and informative results about the nature of bank relationships.

We organize the rest of the paper as follows. Section 2 describes the bank relationship data and details the construction of the firm-specific variables. Section 3 introduces the econometric methodology used in our analysis. Section 4 contains the empirical results, including tests of robustness. Section 5 contains a short discussion and concludes the paper.

2. Data

This section describes the relationship data, presents the distribution of the observed duration of bank relationships, and describes the firm-specific variables used in the cross-sectional regressions. We obtain annual data on bank relationships of firms listed on the Oslo Stock Exchange (OSE) for the years 1979 to 1995 from *Kierulf's Handbook*, an annual publication of the OSE. The handbook contains financial and accounting information on all listed

[2] In a recent paper, Farinha and Santos (2000) also use duration analysis to study bank switching behavior, but emphasize a firm's choice in moving from a single bank relationship to multiple relationships. Consistent with our results, they find that the likelihood of moving to a multiple bank relationship is increasing in relationship duration, firm leverage, and firm growth potential.

firms, as well as other firm-specific information provided to the stock exchange, including firm age and ownership structure. To comply with listing requirements, each firm must report its primary bank relationships, up to a maximum of four. During an average year, 74% of the firms maintain only one bank relationship, while another 17% maintain two relationships, 7% maintain three relationships, and only 2% maintain four or more relationships. We identify a firm as ending a bank relationship when it drops a bank from its list of primary banks or replaces the bank with a new one. We define the duration of a relationship to be the number of consecutive years a firm lists in *Kierulf's Handbook* before it is dropped.

Kierulf's Handbook provides no information about primary bank relationships beyond the identity of each bank. To better understand the nature of these relationships, we collected additional information through phone interviews with 16 randomly chosen firms. Through these interviews, we confirmed the accuracy of the information in the handbook, inquired about the services provided by primary banks, and sought information about other bank relationships held by firms but not listed as primary bank relationships. All firms interviewed use their primary banks for both short-term and long-term borrowing. All of the firms, except for one, also obtain deposit, cash management, and foreign exchange services from their primary bank, and seven of these firms reported purchasing additional derivatives services from their primary bank. We discovered that 11 firms maintain at least one relationship with a bank that is omitted from the list of primary relationships. These firms claim to purchase fewer services from such secondary banks. For ten of the firms, secondary banks provide services related to foreign exchange. In addition, the secondary banks provide lending services to six firms and cash management services to one firm.

Table 1 presents an annual overview of the number of firms and turnover in bank relationships in our sample. We exclude banks from the firm sample because they do not report relationships with other banks. The sample consists of an average of 111 firms per year, representing 95% of all non-bank firms listed on the Oslo Stock Exchange. On average, six firms terminate a relationship every year, while ten firms start a new one. Relationship turnover activity increases during the years 1986 to 1988, a period marked both by substantial financial deregulation and the beginning of a banking crisis in Norway. Deregulation of the banking sector began in 1984 with the removal of interest-rate ceilings and government-mandated lending controls. Subsequently, both bank loan activity and loan losses increased substantially. Kaen and Michalsen (1997) report that the kroner volume of loans increased by 37.5% in 1985 and 23.4% in 1986, while commercial loan losses, as a percentage of commercial bank loans, quadrupled between 1984 and 1987. Concurrent with the increase in the turnover in bank relationships is an increase in the number of firms delisting from the OSE. Many of the delistings

Table 1
Annual overview of sample

This table lists, by year, the total number of firms contained in *Kierulf's Handbook*, the number of Oslo Stock Exchange (OSE) listed banks, the number of firms reporting bank relationships, the number of new bank relationships, the total number of relationships ending, the return on the value-weighted index of all OSE-listed companies, and the number of OSE listings and delistings. A bank relationship in this table is a primary bank relationship, as defined by the firm and reported in *Kierulf's Handbook*. We identify a firm as ending a relationship when it drops a bank from the list or replaces one bank with another. All numbers are from *Kierulf's Handbook*, except the return on the OSE index and the number of listings and delistings, which are from *Oslo Børs Informasjon* AS (OBI). The column containing "Firms reporting bank relationships" is the sample of firms used in our analysis.

Year	OSE firms, including banks	OSE-listed banks	Firms reporting bank relationships	New relationships	Relationship terminations	Return on OSE index	New OSE lists	OSE delists
1979	113	13	96	—	—	86.2%	—	—
1980	109	13	93	5	5	−13.0	6	10
1981	112	14	96	2	1	4.0	5	2
1982	117	14	100	3	4	−16.9	6	1
1983	136	15	115	5	5	92.8	21	2
1984	158	16	140	7	5	27.3	22	0
1985	159	17	138	6	1	27.5	7	6
1986	154	18	133	17	16	−8.8	8	12
1987	143	14	125	14	10	−6.3	5	15
1988	129	13	113	18	12	33.9	3	17
1989	130	15	111	11	6	52.4	11	11
1990	114	11	100	14	7	−14.2	7	24
1991	117	8	106	14	9	−10.0	15	11
1992	121	9	101	16	5	−10.0	11	7
1993	125	9	106	10	4	64.8	15	12
1994	131	7	113	14	5	7.1	17	11
1995	133	6	98	10	6	14.2	20	18
Mean	129.5	12.5	110.8	10.4	6.3	15.3	11.2	9.9

Table 2
Distribution of observed duration of bank relationships

This table lists marginal and cumulative distributions for the observed duration of bank relationships, in years. A bank relationship in this table is a primary bank relationship, as defined by the firm and reported in *Kierulf's Handbook*. We identify a firm as ending a relationship when it drops a bank from the list or replaces one bank with another. The observed duration of a relationship is the number of consecutive years a bank appears in *Kierulf's Handbook* as a primary bank relationship before being dropped. N is the number of firm–bank observations, defined to be the number of firms listing bank relationships multiplied by the number of banks per firm.

Observed duration, in years	Proportion of:					
	All bank relationships ($N=419$)		Relationships formed before 1979 ($N=146$)		Relationships formed between 1979 and 1995 ($N=273$)	
	Marginal	Cumulative	Marginal	Cumulative	Marginal	Cumulative
<1	15.8%	15.8%	17.8%	17.8%	14.7%	14.7%
1	12.6	28.4	2.1	19.9	18.3	33.0
2	11.4	39.8	2.7	22.6	16.1	49.1
3	6.4	46.2	2.7	25.3	8.4	57.5
4	6.2	52.4	2.7	28.0	8.0	65.5
5	7.2	59.6	7.6	35.6	7.0	72.5
6	7.4	67.0	13.0	48.6	4.4	76.9
7	6.4	73.4	9.6	58.2	4.8	81.7
8	5.5	78.9	8.2	66.4	4.0	85.7
9	4.1	83.0	3.4	69.8	4.4	90.1
10	2.9	85.9	4.1	73.9	2.2	92.3
11	3.3	89.2	4.8	78.7	2.6	94.9
12	1.4	90.6	1.4	80.1	1.5	96.4
13	1.7	92.3	0.7	80.8	2.2	98.6
14	0.5	92.8	0.0	80.8	0.7	99.3
15	1	93.8	1.4	82.2	0.7	100.0
⩾16	6.2	100.0	17.8	100.0	0.0	100.0
Mean duration (years)	5.1		7.3		3.9	
Median duration (years)	4.0		6.0		2.0	

during the crisis period are banks, which either merge with another bank or become privately held institutions after being purchased by the Norwegian government.

In the analysis to follow, we define a sample observation as one relationship between a firm and a bank, or a firm–bank relationship. The number of firm–bank observations equals the total number of independent firms in the sample, times the number of relationships maintained by each firm over the sample period. Our data set contains 419 such observations. Table 2 categorizes the firm–bank observations by their observed duration, and gives a sense for the impact of censoring. Column (1) lists the distribution of all bank relationships sorted by the number of years each relationship remains in our sample. Column (2) presents the distribution including only those relationships that begin before 1979. Column (3) excludes the pre-1979 relationships. From column (1), it appears as if only a small proportion of relationships (6.2%) continue beyond the 16-year cutoff of the sample. The median duration of an observed relationship is four years, 40% of the observed relationships do not extend beyond two years, and fully 86% of the firms appear to end their bank relationship by the tenth year. Thus, a first-pass look at the duration data suggests that bank relationships are short-lived and that the end-of-sample cutoff affects only a small proportion of the observations. However, relationships begun before 1979 appear to last relatively longer. Out of these firms, 17.8% continue their bank relationship beyond 1995 and the median relationship lasts six years. Similarly, column (3) indicates that excluding the pre-1979 relationships shifts the weight of the observed durations back to the left. We return to the impact of censoring later.

The firms in our sample maintain relationships with 55 different banks. Of these banks, 14 are international banks, 24 are Norwegian commercial banks, and the remaining 17 are Norwegian savings banks. However, the relationships are concentrated across only a handful of these banks. In an average year, 75% of the firms maintain a relationship with at least one of Norway's two largest commercial banks, Kreditkassen or Den norske Bank. Prior to 1990, there were three major banks, but Bergen Bank and Den norske Credit merged to form Den norske Bank in 1990. Nearly 40% (21) of the banks merged with another bank between 1979 and 1995. Unless otherwise indicated in *Kierulf's Handbook*, we assume customers of merged banks continue with the combined entity, so that no break in the relationship occurs. We explore the sensitivity of our results to market concentration and bank mergers in Section 4.4.

2.1. Firm-specific variables

We identify seven firm-specific characteristics that serve as indicators for a firm's dependence on a bank relationship. Fama (1985), Diamond (1984, 1991), and Rajan (1992) maintain that small and young firms are more likely to suffer from information problems and, therefore, find bank financing useful. Ln Sales, defined to be the natural logarithm of year-end sales deflated by the Norwegian consumer price index (CPI), serves as a proxy for firm size and Age, defined to be the number of years since a firm's founding date, measures firm age. Titman and Wessels (1988) show that highly profitable firms are less dependent on external debt financing, including bank financing. We construct Profitability using the ratio of operating income (sales minus cost of goods sold) to the book value of assets. Holding the level of internal financing constant, firms with more growth opportunities should have greater external financing needs, and therefore be more dependent on bank financing. Tobin's Q is defined as the year-end market value of equity plus book value of debt, divided by the book value of assets. Because much of the value of high-growth firms comes in the form of unrealized investment opportunities, Tobin's Q also serves as a proxy for the level of information asymmetry. Because banks supply such a high proportion of commercial credit in Norway, highly leveraged firms should be more dependent on banks than firms with lower debt levels. Leverage is the book value of debt, divided by the sum of the year-end market value of equity and book value of debt. Firms with multiple bank relationships have more than one potential source of inside bank financing and should therefore face lower switching costs and be less susceptible to holdup threats by any one bank.[3] The dummy variable Multiple Relationships equals one when a firm maintains more than one bank relationship, and zero otherwise. Finally, implicit in the moral hazard models of Boot and Thakor (1994), Rajan (1992), and von Thadden (1995) is the idea that a bank's ability to reduce agency-related problems may be unnecessary if the firm already maintains an effective monitoring mechanism. Ownership Concentration, defined to be the proportion of equity owned by a firm's ten largest shareholders, proxies for an alternative monitoring source.

With the exception of Age, we measure each variable using the values taken from the year prior to either the end of the relationship or the date of right censoring, whichever comes earlier. For the purpose of correctly specifying our regression model, we also define a variable, Age At Start, to be the time elapsed between the firm's founding date and either the start of the relationship, or the date of left censoring, whichever comes last. This definition avoids the spurious correlation that firms become older as duration lengthens. Matching the firm-level and bank relationship data reduces the number of observations from 419 to either 383 or 270, depending on whether we include ownership concentration in the analysis. The initial loss of 7.8% of the observations results primarily from missing accounting data. Ownership Concentration further reduces the

[3] Houston and James (1996) find that investment by multiple-bank firms is less sensitive to fluctuations in internal cash flow than single-bank firms. Detragiache et al. (2000), Ongena and Smith (2000), and Farinha and Santos (2000) study the choice of number of bank relationships.

sample because, for a large part of our sample period, disclosure of ownership information was voluntary and not consistently available.

Table 3 contains summary statistics on the firm-characteristic variables. The average firm in our sample generates sales of 741 million Norwegian Kroner (NOK), or roughly \$100 million, making it larger than the family-owned firms studied by Petersen and Rajan (1994) and Berger and Udell (1995), but smaller than the typical publicly traded U.S. company. A few large firms dominate the OSE, skewing the average size. For example, the average OSE firm had a market value of \$150 million in 1995, placing it in the second smallest size decile of NYSE firms according to Ibbotson (1996). However, the average size of Norway's five largest publicly-listed firms, which account for roughly two-thirds of the total OSE market capitalization, places them into the second largest NYSE size decile. The average firm is 53 years old, earns operating income equal to 1.7% of its assets, and has a market-to-book ratio of 1.4. The average firm also finances 55% of its assets with debt. Although not shown in the table, this is split almost evenly between short-term and long-term liabilities. On average, trade credit (accounts payable plus unearned revenue) represents 9.6% of total liabilities. We cannot observe the amount of bank debt in each firm's capital structure. However, if we assume that our sample firms issue all traded non-bank corporate debt in Norway (commercial paper and corporate bonds), then bank debt represents approximately 80% of the average OSE firm's liabilities. In 1994, total non-financial commercial paper issuances represented 1.6%, and corporate bond issuances 7.1%, of sample firm debt (*Statistical Yearbook of Norway*, *1996*).

Firms maintaining multiple bank relationships generate 44% of the firm–bank observations. There are two reasons for the high incidence of multiple-bank relationships. First, because we count each firm–bank relationship as one observation, a firm with multiple relationships enters as two or more observations. Second, firms with multiple-bank relationships are more likely to end a bank relationship, generating greater turnover in observations. The ten largest shareholders own 69% of the equity of the companies reporting ownership information.

3. Econometric specification

This section develops the econometric methodology employed in analyzing the duration of bank relationships in our sample. We begin by introducing terminology common to duration analysis and then describe the hazard function estimators. Let T represent the duration of time that passes before the occurrence of a certain random event. In the econometrics literature, the passage of time is often referred to as a spell, while the event itself is called a switch. A simple way to describe the behavior of a spell is through its survivor

Table 3

Descriptive statistics for firm-specific characteristics

For this table, we measure each characteristic in the year prior to the end of the bank relationship (or date of right censoring, if censored first). All firm characteristics are from *Kierulf's Handbook* or company annual reports. Sales is year-end sales in millions of Norwegian Kroner, deflated by the Norwegian consumer price index (with 1979 = 100). One US dollar is roughly equivalent to 7 Norwegian Kroner. Ln Sales is the natural logarithm of sales. Age is the time elapsed between the firm's founding date and the year of measurement. Age At Start is the time elapsed between the firm's founding date and either the start of the relationship, or the date of left censoring, whichever comes last. Profitability is the ratio of operating income to book value of assets. Tobin's Q is the ratio of year-end market value of equity plus book value of debt divided by the book value of assets. Leverage is the book value of debt divided by the sum of year-end market value of equity and book value of debt. Multiple Relationships equals one when a firm maintains more than one bank relationship and zero otherwise. Ownership Concentration measures the proportion of equity owned by a firm's ten largest shareholders. Companies report ownership structure information on a voluntary basis; therefore, the number of observations for ownership is lower. A bivariate t-test determines the statistical significance of the correlation coefficients. N is the number of firm–bank observations.

Variable	N	Mean	Median	SD	Minimum	Maximum
Sales	383	741.1	165.8	2,009.9	0.0	30,698.0
Ln Sales	383	7.1	7.4	2.4	0.0	12.6
Age	383	53.0	49.0	43.3	0.0	245.0
Age At Start	383	47.5	44.0	42.1	0.0	243.0
Profitability	383	1.7%	4.7	15.3	−126.2	37.7
Tobin's Q	383	1.4	1.2	0.8	0.3	9.2
Leverage	383	55.4%	58.7	24.9	0.0	98.2
Multiple Relationships	383	44.3%	0.0	49.7	0.0	100.0
Ownership Concentration	270	68.4%	69.5	17.5	4.8	100.0

Correlation matrix	Age	Age At Start	Profitability	Tobin's Q	Leverage	Multiple Relationships	Ownership Concentration
Ln Sales	0.085	0.068	0.303***	−0.209***	0.383***	0.123***	−0.048
Age		0.994	0.121***	0.011	−0.018	0.102***	−0.018
Age At Start			0.119***	0.016	−0.011	0.097**	−0.013
Profitability				−0.075*	−0.071*	0.004	−0.082*
Tobin's Q					−0.543***	−0.102**	−0.009
Leverage						0.173***	−0.010
Multiple Relationships							−0.164***

***Significant at 1%, **significant at 5%, *significant at 10%.

function, $S(t) = \mathrm{P}(T \geqslant t)$, which yields the probability that the spell T lasts at least to time t. The survivor function equals one minus the cumulative distribution function of T. The behavior of a spell can also be described through the use of the hazard function. The hazard function determines the probability that a switch will occur, conditional on the spell surviving through time t, and is defined by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\mathrm{P}(t \leqslant T < t + \Delta t | T \geqslant \mathrm{t})}{\Delta t} = \frac{-\mathrm{d} \log S(t)}{\mathrm{d}t} = \frac{f(t)}{S(t)}, \tag{1}$$

where $f(t)$ is the density function associated with the distribution of spells. Neither the survivor function nor the hazard function provides additional information that could not be derived directly from $f(t)$. Instead, these functions present economically interesting ways of examining the distribution of spells. The hazard function does provide a suitable method for summarizing the relationship between spell length and the likelihood of switching. When $\lambda(t)$ is increasing in t, the hazard function is said to exhibit positive duration dependence, because the probability of ending the spell increases as the spell lengthens. Similarly, negative duration dependence occurs when $\lambda(t)$ is decreasing in t, and constant duration dependence indicates the lack of a relation between $\lambda(t)$ and t.

When estimating hazard functions, it is econometrically convenient to assume a proportional hazard specification, such that

$$\lambda(t, X(t), \beta) = \lim_{\Delta t \to 0} \frac{\mathrm{P}(t \leqslant T < t + \Delta t | T \geqslant t, X(t), \beta)}{\Delta t} = \lambda_0(t) \exp{(\beta' X_t)}, \tag{2}$$

where X_t is a set of observable, possibly time-varying explanatory variables, β is a vector of unknown parameters associated with the explanatory variables, $\lambda_0(t)$ is the baseline hazard function and $\exp(\beta' X_t)$ is chosen because it is non-negative and yields an appealing interpretation for the coefficients, β. The logarithm of $\lambda_0(t, X(t), \beta)$ is linear in X_t. Therefore, β reflects the partial impact of each variable in X on the log of the estimated hazard rate.

The baseline hazard $\lambda_0(t)$ determines the shape of the hazard function with respect to time. Eq. (2) can be estimated without specifying a functional form for the baseline hazard. The Cox (1972) partial likelihood model bases estimation of β on the ordering of the duration spells. Because it specifies no shape for $\lambda_0(t)$, we refer to the Cox partial likelihood model as "semi-parametric." Two commonly used parametric specifications for the baseline hazard are the Weibull and the exponential distributions. The Weibull specification assumes $\lambda_{\mathrm{o}}(t) = \lambda \alpha t^{\alpha - 1}$, and allows for duration dependence. When $\alpha > 1\,(\alpha < 1)$, the distribution exhibits positive (negative) duration dependence. The exponential distribution, which exhibits constant duration dependence, is nested within the Weibull as the case $\alpha = 1$. Using maximum likelihood methods, we estimate hazard functions using the Cox (1972) partial likelihood model, Weibull specification, and exponential specification.

We assume that the explanatory variables in our model can vary through time. To obtain interpretable estimates from the proportional hazard models, we require that our variables be either "defined" or "ancillary" with respect to the duration of a spell (see Kalbfleisch and Prentice, 1980). A defined variable follows a deterministic path. Age is an example of a defined variable because its path is set in advance of the bank relationship, and varies deterministically with relationship duration. An ancillary variable has a stochastic path, but the path cannot be influenced by the duration of the spell. For the reported results, we also assume that the conditional likelihood of ending a spell depend only on the value of an ancillary variable at time t, and not on past or future realizations of the variable. We later discuss the sensitivity of our results to these assumptions.

3.1. Censoring

Censoring affects a total of 76% of the original 419 observations. Not knowing when a relationship starts, or when it ends, or both, means we are unable to observe the true duration of the relationship for these observations. With no adjustment to account for censoring, maximum likelihood estimation of the proportional hazard models produces biased and inconsistent estimates of model parameters.

Accounting for right-censored observations can be accomplished by expressing the log-likelihood function as a weighted average of the sample density of completed duration spells and the survivor function of uncompleted spells (see Kiefer, 1988). Controlling for left censoring is less straightforward and often ignored in economic duration analysis. However, Heckman and Singer argue that the biases induced by left censoring can be as severe as those created by right censoring. To analyze the sensitivity of our results to left censoring, we incorporate an estimation strategy shown by Heckman and Singer (1984) to yield consistent, albeit inefficient, estimates of duration under left censoring.

4. Empirical results and robustness tests

Before turning to cross-sectional estimates of the hazard function, we first examine some simple plots of the distribution of relationship spells. Fig. 1 shows estimates of the survivor function, $S(t) = \mathrm{P}(T \geqslant t)$. The plots illustrate the impact of right censoring on estimates of relationship duration. The lighter solid line in Fig. 1 traces the survivor function without adjusting for right censoring. The darker solid line includes the adjustment. To construct the

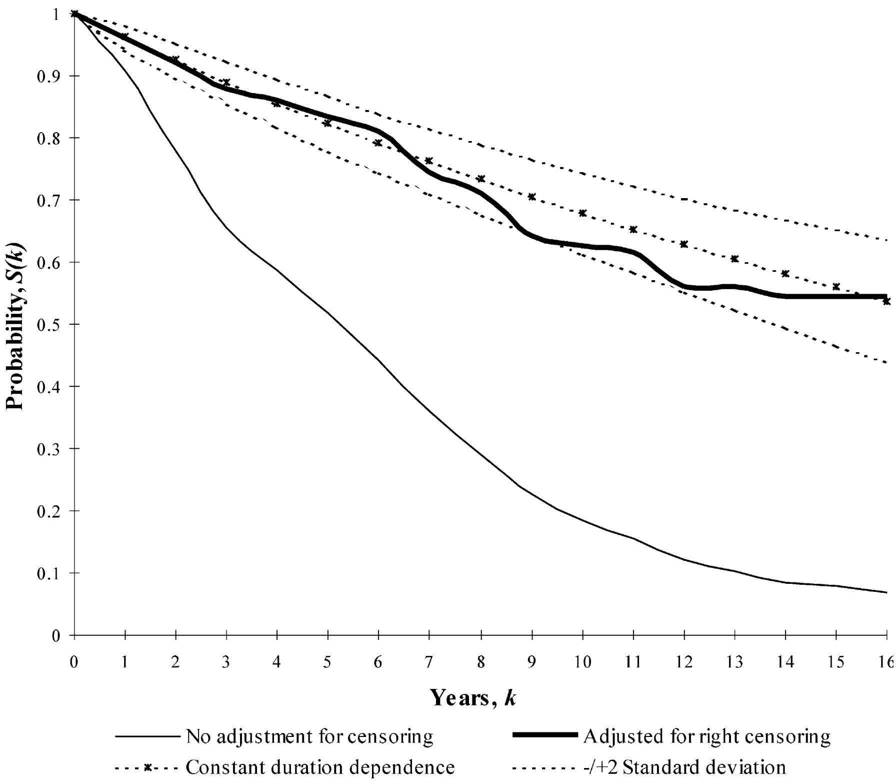


Fig. 1. Non-parametrically estimated survivor functions with and without adjustment for right censoring. The figure is based on estimates of the survivor function $\hat{S}(k) = \prod_{i=0}^{k}(1 - \hat{\lambda}(i))$, where $\hat{\lambda}(i)$ is the sample estimator for the probability a firm ends a bank relationship, conditional on the relationship lasting i periods. The heavy-dashed line plots the implied censored-robust estimate for a survivor function calculated under the restriction of constant duration dependence. The two lighter-dashed lines plot approximate 95% confidence intervals around the restricted estimate. Number of firm–bank observations: 383.

survivor function, we use the Kaplan and Meier (1958) estimator, $\hat{S}(k) = \prod_{i=0}^{1}(1 - \hat{\lambda}(i))$, where $\hat{S}(k)$ is the estimated probability that a relationship survives beyond year k. Without a correction for right censoring, $\hat{\lambda}(i)$ is the number of firms leaving the data set during year i, divided by the total number of firms remaining in the data set at the beginning of year i. With a correction for censoring, $\hat{\lambda}(i)$ is the number of firms observed to terminate a relationship in year i, divided by the number of surviving firms.

The survivor function decreases quickly when the censoring adjustment is omitted. The estimated likelihood of a relationship surviving beyond the beginning of its fifth is only 48% (see Table 2), while the estimated chance of surviving past the beginning of the 16th year is 6.2%. In contrast, when the adjustment for right censoring is added to the estimation equation, the likelihood of surviving past the beginning of year five increases to 83%, and the likelihood of surviving beyond the 16th year increases to 54%. Without an adjustment for right censoring, relationships appear to end relatively quickly, biasing inferences on the durability of relationships.

Fig. 1 also depicts a test for duration dependence. The dashed line with asterisks plots the survivor function adjusted to account for right censoring, assuming constant duration dependence. The two lighter dashed lines define a confidence interval that incorporates approximately 95% of the expected values surrounding the estimate restricted to constant duration dependence. Because the thick solid line lies within the 95% confidence interval bands, the nonparametric estimates cannot reject the hypothesis of constant duration dependence. In other words, the plots suggest that the likelihood of ending a relationship is no greater for long-term relationships than for short-term relationships.

4.1. Estimates using the partial likelihood model

Starting in Table 4, we estimate the impact of the firm-specific characteristics on the conditional probability of ending a bank relationship using the proportional hazard specification in Eq. (2) and the semi-parametric Cox (1972) partial likelihood model. There are two ways to interpret the signs on the slope estimates represented by β in Eq. (2). First, each estimate represents the partial impact of a firm-specific characteristic on the probability of ending a relationship, holding duration constant. Second, because duration is inversely related to the hazard rate, a positive (negative) coefficient estimate implies a shorter (longer) duration.

We present four different models in Table 4. The first model uses all of the firm-specific variables, but does not correct for right censoring. It shows that firms maintain shorter relationships when they are smaller, younger, have more growth opportunities, and are highly leveraged. Model 2 includes the adjustment for right-censored observations. After applying the adjustment, the coefficients on ln Sales and Leverage nearly double in value, the estimate associated with Multiple Relationships jumps ten-fold and becomes significant at the 5% level, and Age At Start and Tobin's Q are no longer statistically significant. Moreover, the difference in the log-likelihood values between Models 1 and 2 rejects the hypothesis that right censoring has no impact on the coefficient estimates at the 1% level. Models 3 and 4 adjust for right censoring but drop Ownership Concentration, allowing us to increase the sample size by 42%. Excluding Ownership Concentration and including the additional observations does not alter the magnitude of the relations implied by Model 2. However, Profitability becomes statistically significant at a 5% level in Model 4. The positive estimate implies that profitable firms end bank

Table 4
Partial likelihood estimates of proportional hazard model

The estimates in this table are based on ML estimation of the proportional hazard model using the Cox (1972) partial likelihood function. The coefficients measure the partial impact of each variable on the likelihood a relationship terminates, conditional on duration. Ln Sales is the natural logarithm of year-end sales measured in 1979 Norwegian Kroner. Age At Start is the time elapsed between the firm's founding date and either the start of the relationship, or the date of left censoring, whichever comes last. Profitability is the ratio of operating income to book value of assets. Tobin's Q is the ratio of year-end market value of equity plus book value of debt divided by the book value of assets. Leverage is the book value of debt divided by the sum of year-end market value of equity and book value of debt. Multiple Relationships equals one when a firm maintains more than one bank relationship and zero otherwise. Ownership Concentration measures the proportion of a firm's equity owned by its ten largest shareholders. With the exception of Age At Start, we measure all variables in the year prior to termination or right censoring, whichever comes later. The estimates in Model 1 are not adjusted for right censoring. The estimates in Models 2–4 are adjusted for right censoring. Coefficients are listed on the first row in each cell, with standard errors reported below in parentheses. The last column lists the value of the optimized log likelihood function $L(\theta)$. N is the number of firm–bank observations.

Model	N	Ln Sales	Age At Start	Profitability	Tobin's Q	Leverage	Multiple Relationships	Ownership Concentration
1	270	−0.096***	−0.00541***	0.392	0.184**	1.066***	0.058	0.300
		(0.032)	(0.00160)	(0.541)	(0.092)	(0.380)	(0.131)	(0.360)
2	270	−0.215***	−0.00238	1.588	0.244	2.143***	0.608***	0.666
		(0.058)	(0.00280)	(1.161)	(0.174)	(0.708)	(0.254)	(0.681)
3	270	−0.213***	−0.00227	1.567	0.249	2.161***	0.559***	—
		(0.058)	(0.00278)	(1.181)	(0.166)	(0.704)	(0.248)	
4	383	−0.195***	−0.00350	1.741**	0.237	2.088***	0.520***	—
		(0.042)	(0.00241)	(0.882)	(0.146)	(0.515)	(0.191)	

***Significant at 1%, **significant 5%, *significant at 10%.

relationships earlier, consistent with the argument that profitable firms are less dependent on bank financing and therefore less susceptible to holdup costs. Moreover, because credit denials and forced terminations by banks are likely reserved for unprofitable firms, the significantly positive estimate produced in Model 4 suggests that banks do not initiate the terminations that we observe.

The corrected estimates suggest that the conditional likelihood of ending a bank relationship decreases in firm size, and increases in firm leverage, profitability, and when firms maintain multiple-bank relationships. In other words, small, highly leveraged, and profitable firms, and firms with more than one bank relationship tend to maintain shorter relationships. This basic pattern, with some alterations, will persist throughout the rest of this analysis.

4.2. Results using restricted baseline hazard models

To check the sensitivity of the regression results to restrictions on the baseline hazard rate, Table 5 reports the results from the exponential and Weibull specifications of the baseline hazard function presented in Eq. (2). For each specification, we report the results of three regressions. The first regression includes Ownership Concentration, the second regression drops Ownership Concentration but uses the same 270 observations as in the first regression, and the third regression drops Ownership Concentration and increases the sample size to 383. All of the regressions in Table 5 include an adjustment for right censoring.

The magnitude and significance of the coefficient estimates are similar to those in Table 4, indicating that they are not sensitive to the specification of the baseline hazard function. Profitability continues to be positive and statistically significant over all specifications. In addition, the negative coefficient estimate on Age At Start is significant at the 10% level in Model 6.

The Weibull model estimates α to be significantly greater than one, implying that bank relationships exhibit positive duration dependence. This result contrasts with our unconditional estimates in Fig. 1 that found no duration dependence. Heckman and Singer (1984) provide an explanation for these conflicting results. They argue that heterogeneity across observations biases non-parametric estimates in favor of negative duration dependence. By controlling for heterogeneity through the firm-specific characteristics we mitigate this bias and uncover positive duration dependence. Although we do not report the results, we also estimate the conditional hazard function assuming the baseline hazard follows a log–logistic distribution. The log–logistic allows for non-monotonic duration dependence. The estimates suggest that the likelihood of ending a bank relationship increases over time in the earlier part of a relationship (positive duration dependence), but decreases later in the relationship (negative duration dependence). Plots of the estimated survivor function using partial likelihood estimates produce a similar shape.

Table 5
Parametric estimation of proportional hazard model

The estimates in this table are based on ML estimation of the proportional hazard model using the exponential and the Weibull distributions as baseline hazard rates. The parameter α measures the degree of duration dependence. The exponential model assumes $\alpha = 1$. Ln Sales is the natural logarithm of year-end sales measured in 1979 Norwegian Kroner. Age At Start is the time elapsed between the firm's founding date and either the start of the relationship, or the date of left censoring, whichever comes last. Profitability is the ratio of operating income to book value of assets. Tobin's Q is the ratio of year-end market value of equity plus book value of debt divided by the book value of assets. Leverage is the book value of debt divided by the sum of year-end market value of equity and book value of debt. Multiple Relationships equals one when a firm maintains more than one bank relationship and zero otherwise. Ownership Concentration measures the proportion of a firm's equity owned by its ten largest shareholders. With the exception of Age At Start, we measure all variables in the year prior to termination or right censoring, whichever comes later. All estimates are adjusted for right censoring. Coefficients are listed on the first row in each cell, with standard errors reported below in parentheses. N is the number of firm–bank observations.

Model	N	$\hat{\alpha}$	$\hat{\beta}$								Median duration
			Intercept	Ln Sales	Age At Start	Profitability	Tobin's Q	Leverage	Multiple Relationships	Ownership Concentration	
Exponential	270	1	−4.275***	−0.241***	−0.00396	1.911*	0.354	2.553***	0.757***	0.748	22.1
(1)		—	(0.878)	(0.070)	(0.00302)	(1.115)	(0.240)	(0.843)	(0.283)	(0.838)	(3.1)
Exponential	270	1	−3.771**	−0.239***	−0.00382	1.918*	0.357*	2.577***	0.702**	—	21.9
(2)		—	(0.280)	(0.069)	(0.00298)	(1.138)	(0.212)	(0.823)	(0.280)		(3.0)
Exponential	383	1	−3.601***	−0.218***	−0.00352	2.124***	0.268	2.281***	0.659***	—	19.8
(3)		—	(0.561)	(0.053)	(0.00259)	(0.997)	(0.195)	(0.628)	(0.231)		(2.3)
Weibull	270	1.260^{+}	−3.906***	−0.202***	−0.00404	1.634*	0.314	2.223***	0.611***	0.627	18.3
(4)		(0.162)	(0.730)	(0.055)	(0.00236)	(0.901)	(0.194)	(0.661)	(0.229)	(0.664)	(3.0)
Weibull	270	1.257^{+}	−3.488***	−0.201***	−0.00391	1.647*	0.318*	2.249***	0.566**	—	18.2
(5)		(0.162)	(0.504)	(0.054)	(0.00233)	(0.892)	(0.172)	(0.646)	(0.224)		(2.9)
Weibull	383	1.351^{+++}	−3.260***	−0.178***	−0.00344*	1.752**	0.238*	1.933***	0.491***	—	15.8
(6)		(0.135)	(0.408)	(0.038)	(0.00183)	(0.717)	(0.141)	(0.444)	(0.168)		(1.9)

$^{+++}$, $^{++}$, $^{+}$ $\alpha = 1$ can be rejected at 1%, 5%, and 10%, respectively. ***Significant at 1%, **significant at 5%, *significant at 10%.

The coefficient estimates on the explanatory variables from the log–logistic regressions are nonetheless very close to those reported in Tables 4 and 5.

The parametric models allow us to assign economic meaning to the censored-adjusted estimates in Table 5. The estimated median duration is given by $t = -\ln(0.5)^{1/\hat{\alpha}}/\exp(\hat{\beta}'\bar{X})$. In Table 5, estimates of the median duration of a bank relationship range between 15.8 and 22.1 years, compared with a median of four years based on the unadjusted observations in column (1) of Table 2. Fitted hazard rates illustrate the magnitude of positive duration dependence. For instance, Model 6 implies that the likelihood of ending a relationship over the subsequent year after surviving one year is 1.7%, after two years 2.6%, after five years 3.9%, after ten years 5.0%, and after 20 years 6.5%.

We can also quantify the estimated impact of the explanatory variables on the expected duration of the relationship. Model 6 implies that the relationship of a single-bank firm lasts 7.7 years longer than a similar firm with multiple bank relationships. A 50% decrease in firm size from its median (to NOK 83 million) reduces the median fitted duration by 1.9 years, a 50% increase above the median value of leverage (to 83%) reduces the duration by 6.6 years, and an increase in profitability from its mean of 1.7% to its median of 4.7% reduces duration by 0.81 years. Taking these changes jointly, a sample firm characterized as small, highly leveraged, and profitable maintains a relationship for only 8.6 years, compared with 15.8 years for the average firm in the sample. By contrast a firm that is 50% larger than the median (NOK 249 million), carries 50% less debt than average (28%), and has a return on assets of 1.0% maintains a relationship for a median of 30.9 years.

4.3. Robustness tests

Table 6 checks the robustness of the results in Tables 4 and 5. We analyze previously unexamined reasons for observing bank terminations, apply corrections for left censoring, and evaluate the assumed process governing the path of the explanatory variables. For all of the results, we report only the Weibull specification without Ownership Concentration.

4.3.1. Other explanations for terminations

Norwegian bank regulations prohibit a bank from lending more than 5% of its equity capital to any one firm. In Model 1 of Table 6, we add the variable Bank Size as a measure of a firm's debt level relative to the loan capacity of its banks. Our aim is to determine whether firms terminate relationships because they outgrow the lending capacity of their banks. To construct Bank Size, we first obtain the book value of equity for Norway's 15 largest banks. For all other banks, we assume the book value of equity is zero. This assumption biases downward the true lending capacity of the banks. We then define Bank Size to be ln[1 + (book value of firm debt)] −ln[1 + (book value of equity of the

Table 6
Estimation of Weibull proportional hazard hazard model: robustness tests

The estimates in this table are based on ML estimation of the proportional hazard model using the Weibull distribution as the baseline hazard rate. Profitability is the ratio of operating income to book value of assets. Bank Size is ln (1 + book value of firm debt)−ln(1 + book value of equity of the firm's bank(s)). If the firm maintains more than one relationship, the denominator is the sum of the book value of equity across the banks. Large Bank is a dummy variable taking the value of one when a relationship is with one of Norway's two largest banks, Den norske Bank or Kreditkassen. Merger is a dummy variable set equal to one when a relationship bank was involved in a merger in the year prior to, or concurrent with, the termination of a relationship. Replace is a dummy variable that equals one if a bank is replaced by another bank upon termination. Distress is a dummy variable that equals one if a firm's bank becomes distressed in the year prior to, or concurrent with, the termination of a relationship. Relative takes the value of one if the duration of the terminated relationship is greater than the average duration of the other relationships maintained at the point of termination (single-bank relationships are excluded). Tobin's Q is the ratio of year-end market value of equity plus book value of debt divided by the book value of assets. Leverage is the book value of debt divided by the sum of year-end market value of equity and book value of debt. Age At Start is the time elapsed between the firm's founding date and either the start of the relationship, or the date of left censoring, whichever comes last. Multiple relationships equals one when a firm maintains more than one bank relationship and zero otherwise. Ln Sales is the natural logarithm of year-end sales measured in 1979 Norwegian Kroner. With the exception of Age At Start, we measure all firm-specific variables in the year prior to termination or right censoring, whichever comes later. In the Start 1985 model, the starting date of the sample is assumed to be 1985. The No Left Censoring model excludes left-censored observations from estimation. The All Lagged Model measures all of the explanatory variables (with the exception of Age At Start, which continues to be measured at the beginning of the relationship) two years prior to a relationship end, or upon right censoring. Coefficients are listed on the first row in each cell with standard errors reported below in parentheses. N is the number of firm–bank observations.

Model	N	$\hat{\alpha}$	$\hat{\beta}$								Median duration
			Intercept	Ln Sales	Age At Start	Profitability	Tobin's Q	Leverage	Multiple Relationships	Added variable	
Bank Size	383	1.350^{+++}	−3.265***	−0.175***	−0.00346*	1.752**	0.230	1.929***	0.481***	0.00464	15.9
(1)		(0.136)	(0.481)	(0.039)	(0.00184)	(0.720)	(0.142)	(0.445)	(0.173)	(0.01510)	(1.9)
Large Bank	383	1.364^{+++}	−2.936***	−0.159***	−0.00289	1.584**	0.213	1.784***	0.354**	−0.403**	15.6
(2)		(0.136)	(0.391)	(0.038)	(0.00179)	(0.700)	(0.134)	(0.437)	(0.170)	(0.165)	(1.9)
Merger	383	1.351^{+++}	−3.257***	−0.178***	−0.00346*	1.773**	0.234	1.932***	0.496***	0.058	15.8
(3)		(0.135)	(0.410)	(0.038)	(0.00183)	(0.725)	(0.143)	(0.447)	(0.169)	(0.289)	(1.9)

Distress	383	1.351^{+++}	−3.298***	−0.175***	−0.00346*	1.842**	0.220	1.911***	0.541***	0.405	16.0
(4)		(0.136)	(0.425)	(0.038)	(0.00185)	(0.737)	(0.148)	(0.451)	(0.177)	(0.306)	(2.0)
Replace	383	1.452^{+++}	−3.756***	−0.142***	−0.00416**	1.499**	0.307**	1.511***	0.560***	1.509***	19.0
(5)		(0.136)	(0.455)	(0.038)	(0.00119)	(0.652)	(0.145)	(0.481)	(0.189)	(0.204)	(3.0)
Relative	170	2.054^{+++}	−2.468***	−0.121***	−0.00165	1.315*	0.340**	2.274***	—	−1.252***	9.8
(6)		(0.233)	(0.432)	(0.028)	(0.00145)	(0.672)	(0.155)	(0.418)		(0.136)	(0.7)
Start 1985	355	1.214^{+}	−3.013***	−0.232***	−0.00310	2.304***	0.349***	2.302***	0.571***		12.0
(7)		(0.163)	(0.413)	(0.049)	(0.00234)	(0.937)	(0.116)	(0.532)	(0.194)	—	(1.9)
No Left	86	1.485^{+}	−1.970**	−0.105	−0.00474	−0.226	0.028	0.946	0.430		6.7
(8)		(0.308)	(0.838)	(0.064)	(0.00330)	(1.385)	(0.360)	(0.798)	(0.334)	—	(1.2)
All Lagged	345	1.625^{+++}	−3.032***	−0.113***	−0.00132	−0.808	−0.049	1.300***	0.376**		16.4
(9)		(0.158)	(0.386)	(0.026)	(0.00197)	(0.701)	(0.155)	(0.332)	(0.155)	—	(1.7)

[a] $^{+++}$, $^{++}$, $^{+}\alpha=1$ can be rejected at 1%, 5%, and 10% respectively. ***Singificant at 1%, **significant at 5%, *significant at 10%.

firm's banks)], where if the firm maintains more than one relationship, the second term includes the sum of the book value of equity across the banks. The estimate for this variable (appearing in the column labeled "added variable") is statistically insignificant, indicating that bank lending limits are not binding at the time a firm terminates a relationship.

For the second model presented in Table 6, we identify all relationships before 1990 involving Bergen Bank, Den norske Credit, or Kreditkassen, and after 1990 involving either Den norske Bank or Kreditkassen, with the dummy variable Large Bank. The large bank estimate in the added variable column is statistically significant. Holding the other variables constant at their mean values, the estimate implies that relationships with Den norske Bank and Kreditkassen last a median of 18 years, while relationships with other banks last a median of 12 years. Despite the significance of Large Bank, all of the original regression results remain intact. The impact of Large Bank could suggest that firms with relatively large or complex financial demands have little choice but to rely on one of Norway's two largest banks. That is, large firms become locked into large banks because no other large bank alternatives exist. Consistent with this conjecture, we find in results not reported here that firms tend to switch from smaller banks to larger banks. But this explanation runs counter to the fact that many of our firms also maintain relationships with large international banks like Barclays, Citibank, Chase Manhattan, Chemical Bank, and Credit Lyonnaise. Alternatively, the largest banks in Norway could be the best quality banks and the Large Bank result could imply that firms value long relationships with high quality banks. Billett et al. (1995) argue that the value of a bank relationship is positively related to the quality of the bank. According to a survey reported in Ongena and Smith (1998), Norwegian treasury managers perceive Den norske Bank and Kreditkassen to be the highest quality banks in Norway.

A merger between two banks can interrupt the benefits flowing from a bank relationship, culminating in the termination of a relationship. Sapienza (1998) finds the likelihood of terminating a lending relationship is greater after a bank merger, particularly when the borrower is a small customer of the bank targeted in the acquisition. Her study does not, however, control for relationship duration. To study the impact of bank mergers on our hazard rate estimates, we add a dummy variable Merger, defined to be one when a bank merger occurs in either the year of, or the year prior to, a termination. Model 3 of Table 6 indicates that Merger has no impact on the likelihood of terminating a relationship, or on the other coefficient estimates in the regression. Though not reported here, we run also ran a regression specification similar to Sapienza's specification, controlling for the influence of duration. In accordance with her variable definition, we assign a value of one to the merger variable whenever a bank merger occurs at any point during the course of the relationship. In contrast to the results in Sapienza, we find that firms are less likely to end a relationship when a merger has occurred at some point in the past. Throughout all regressions, the relation between the hazard rate and our firm-specific characteristics remains unchanged.

Financial distress at banks could interrupt the flow of relationship benefits and induce firms (or banks) to terminate relationships early. Model 4 in Table 6 adds a dummy variable Distress that equals one if a firm's bank announces insolvency in the year of, or prior to, termination. The bank distress data come from Ongena, Smith et al. (2000). The results in the added variable column indicate that the event of bank distress does not influence relationship termination rates.

Because the decision to switch banks may be different from the decision to simply drop a bank, we include a dummy variable Replace in Model 5 that takes the value of one when a firm, in the termination year of an incumbent relationship, reports a new relationship. By using this ex post information, we assume that a firm knew of a viable replacement bank before the termination decision was made. The coefficient on Replace is positive, statistically different from zero, and economically significant. The estimate implies that firms with viable replacement banks end relationships after roughly five years, while otherwise identical firms without a replacement bank end their relationships after 24 years. Therefore, a firm that finds it easy to replace its incumbent relationship with a competitor reduces considerably the time it spends in a given relationship. Holding all else constant, this result suggests that firms without viable alternatives could be locked in to one relationship. However, the result could also imply that firms satisfied with a long-term relationship do not seek out other sources of financing. In any case, the signs and magnitude of the original firm-specific estimates remain robust to the addition of Replace.

In Model 6 we seek to determine whether multiple-bank firms terminate older relationships in favor of new ones, or terminate a series of short bank relationships while maintaining one long relationship. We do this by restricting the regression sample to include only multiple-bank firms, and adding the dummy variable Relative to the regression specification. Relative takes the value of one when the relationship being terminated is longer than the average length of the firm's other relationships.[4] The estimate associated with Relative is negative and statistically significant, suggesting that multiple-bank firms tend to drop and replace their newer relationships in favor of keeping one long relationship. According to the estimate, multiple-bank firms turn over a new relationship after four years, compared with 15 years for their long-term relationships. This result helps to explain why we see relatively long pre-1979 relationships in Table 2. Many of the terminations observed during the sample

[4] For example, suppose a firm maintains a relationship with both Bank A and Bank B, and we observe a termination with Bank A. Then, Relative equals one if the relationship between the firm and Bank A is longer than the relationship with Bank B.

period are new relationships maintained by multiple-bank firms. When we interact Relative with In Sales (not reported), we find that small firms are even more likely to turnover new relationships and keep a long-term relationship. The finding on Relative is important because it suggests that firms with viable alternatives to single-bank financing find value in a long-term relationship. In other words, freed from holdup problems, firms–particularly small ones–prefer long-term bank relationships.

4.3.2. Left censoring

Up to now, we have only corrected for right censoring and ignored the impact of left censoring. In Models 7 and 8, we conduct two experiments to examine the sensitivity of our results to the presence of left censoring. In Model 7, we assume the first observed year in the sample is 1985 rather than 1979. If our regression results are sensitive to left censoring, then changing the start date should induce instabilities in the parameter estimates. The estimates are robust to the 1985 starting date. They are also robust to a variety of other start dates not reported in the table. In Model 8, we eliminate all left censored observations. Heckman and Singer (1984) show that this strategy yields consistent though inefficient estimates of duration in the presence of left censoring. The median estimate of duration for Model 8 is 6.7 years. This estimate is larger than the unadjusted median estimate of four years in Table 2, but significantly smaller than the estimates that reflect adjustments only for right censoring. It is likely that correcting for right censoring without also correcting for left censoring overestimates the median duration of a relationship. With only 86 observations, none of the coefficient estimates in Model 8 is statistically significant. However, with the exception of Profitability, all of the parameter estimates maintain the same sign as in the earlier results.

4.3.3. Process governing explanatory variables

Lastly, we return to the restrictions imposed in Section 3.3 on the stochastic process assumed to govern the firm-specific explanatory variables. In particular, we check the sensitivity of the results to the assumption that the hazard rate at date t is independent of the path followed by the firm characteristic variables prior to date t. We first change the dating of the variables and re-estimate the model. Model 9 of Table 6, labeled "All-Lagged", lags all of the explanatory variables so that they are measured two years prior to termination or censoring (with the exception of age at start, which continues to be measured at the beginning of the relationship). This modification results in little change to the original results. Next, we examine a model that explicitly allows for some path dependence in the firm characteristic variables. We do this by allowing each variable to vary through the duration of the relationship between a high and low cohort, while holding the other variables constant with respect to time. Though we do not report the results, this specification does not alter our earlier results.

5. Discussion and conclusion

This paper presents new evidence on the value of long-term bank relationships by examining the duration of relationships. We document the presence of positive duration dependence in bank relationships. In other words, firms become more likely to end a bank relationship as a relationship matures. Taken alone, this result suggests that the value of relationships decline through time, and that firms are able to end relationships early, possibly to avoid lock-in. This inference is strengthened by the fact that small, young, and highly leveraged firms maintain the shortest relationships. Although theory suggests that such bank-dependent firms are the most susceptible to lock-in, our findings imply that switching costs are low enough to permit these firms to change banks often.

Holding other firm characteristics constant, we show that competing bank relationships reduce the market power of any one bank, making long-term relationships more valuable. Although firms with multiple bank relationships terminate relationships frequently, they do so by terminating newer relationships and keeping long-term ones. Intuitively, the existence of alternative sources of bank credit reduces the ability for any one bank to threaten holdup. With lower holdup costs, a long-term relationship becomes more valuable to the multiple-bank firms.

We also find some indication that firms terminate relationships as they outgrow their banks. Firms tend to switch from small banks to larger banks, and maintain the longest relationships with Norway's two largest banks. However, we find no evidence that this preference for larger banks arises as a result of limited capacity at other banks. Instead, growing firms could prefer the higher quality services offered by the large Norwegian banks.

The evidence presented here should be useful to future theorists interested in modeling the value of bank relationships. However, one should take caution in drawing far-reaching conclusions from our study. Our data reveals very little about the actual nature of the relationships. We are unable to observe how the price and quantity of lending change over the course of the relationship and do not know the other types of banking services offered to customers in a bank relationship. Indeed, an ideal extension of this study would be to obtain a time-series of relationship-specific information about banks and their customers and examine the duration of the relationship as a function of relationship-specific variables.

References

Berger, A.N., Udell, G.F., 1995. Relationship lending and lines of credit in small firm finance. Journal of Business 68, 351–381.

Berglöf, E., von Thadden, E.L., 1994. Short-term versus long-term interests: capital structure with multiple investors. Quarterly Journal of Economics 109, 1055–1084.

Best, R., Zhang, H., 1993. Alternative information sources and the information content of bank loans. Journal of Finance 48, 1507–1522.

Billett, M.T., Flannery, M.J., Garfinkel, J.A., 1995. The effect of lender identity on a borrowing firm's equity return. Journal of Finance 50, 699–718.

Boot, A.W.A., Thakor, A.V., 1994. Moral hazard and secured lending in an infinitely repeated credit market game. International Economic Review 35, 899–920.

Chemmanur, T.J., Fulghieri, P., 1994. Reputation, renegotiation and the choice between bank loans and publicly traded debt. Review of Financial Studies 7, 475–506.

Cox, D., 1972. Regression models and life tables. Journal of the Royal Statistical Society 24, 187–201.

Detragiache, E., Garella, P., Guiso, L., 2000. Multiple versus single banking relationships: theory and evidence. Journal of Finance 55, 1133–1161.

Diamond, D., 1984. Financial intermediation and delegated monitoring. Review of Economic Studies 51, 393–414.

Diamond, D., 1991. Monitoring and reputation: the choice between bank loans and privately placed debt. Journal of Political Economy 99, 689–721.

Fama, E.F., 1985. What's different about banks? Journal of Monetary Economics 15, 5–29.

Farinha, L.A., Santos, J.A.C., 2000. Switching from single to multiple bank lending relationships: determinants and implications. Unpublished Working Paper, Bank for International Settlements, Basle.

Greenbaum, S.I., Kanatas, G., Venezia, I., 1989. Equilibrium loan price under the bank–client relationship. Journal of Banking and Finance 13, 221–235.

Heckman, J.J., Singer, B., 1984. Econometric duration analysis. Journal of Econometrics 24, 63–132.

Hoshi, T., Kashyap, A., Scharfstein, D., 1991. Corporate structure, liquidity and investment: evidence from Japanese industrial groups. Quarterly Journal of Economics 106, 33–60.

Houston, J., James, C., 1996. Bank information monopolies and the mix of private and public debt claims. Journal of Finance 51, 1863–1889.

Ibbotson, R., 1996. Stocks, Bonds, Bills, and Inflation Yearbook. Ibbotson Associates, Chicago.

James, C., 1987. Some evidence on the uniqueness of bank loans. Journal of Financial Economics 19, 217–235.

James, C., Wier, P., 1990. Borrowing relationships, intermediation and the cost of issuing public securities. Journal of Financial Economics 28, 149–171.

Kaen, F.R., Michalsen, D., 1997. The effects of the Norwegian banking crisis on Norwegian bank and non-bank stocks. Journal of Multinational Financial Management 7, 83–111.

Kalbfleisch, J.D., Prentice, R.L., 1980. The Statistical Analysis of Failure Time Data. Wiley, New York.

Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53, 457–481.

Kiefer, N.M., 1988. Economic duration data and hazard functions. Journal of Economic Literature 26, 646–679.

Lummer, S.L., McConnell, J.J., 1989. Further evidence on the bank lending process and the capital market response to bank loan agreements. Journal of Financial Economics 25, 99–122.

Mikkelson, W.H., Partch, M.M., 1986. Valuation effects of security offerings and the issuance process. Journal of Financial Economics 15, 31–60.

Nilsen, R., 1995. Markedet for risikokapital. In: Vale, B. (Ed.), Norske Finansmarkeder, Norsk Penge-og Valutapolitikk. Norges Bank, Oslo, pp. 104–109.

Ongena, S., Smith, D.C., 1998. Quality and duration of bank relationships. In: Birks, D.F. (Ed.), Global Cash Management in Europe. Macmillan, London, pp. 224–235.

Ongena, S., Smith, D.C., 2000. What determines the number of relationships? Cross country evidence. Journal of Financial Intermediation 9, 26–56.

Ongena, S., Smith, D.C., Michalsen, D., 2000. Firms and their distressed banks: lessons from the Norwegian banking crisis (1987–1991). Unpublished working paper. Federal Reserve Board.

Petersen, M.A., Rajan, R.G., 1994. The benefits of lending relationships: evidence from small business data. Journal of Finance 49, 3–37.

Petersen, M.A., Rajan, R.G., 1995. The effect of credit market competition on lending relationships. Quarterly Journal of Economics 110, 407–443.

Preece, D., Mullineaux, D., 1994. Monitoring of financial intermediaries: banks vs. nonbanks. Journal of Financial Services Research 8, 193–202.

Rajan, R.G., 1992. Insiders and outsiders: the choice between informed and arm's-length debt. Journal of Finance 47, 1367–1400.

Sapienza, P., 1998. The effects of banking mergers on loan contracts. Unpublished working paper, Northwestern University.

Sharpe, S.A., 1990. Asymmetric information, bank lending and implicit contracts: a stylized model of customer relationships. Journal of Finance 45, 1069–1087.

Shockley, R., Thakor, A.V., 1998. Bank loan commitment contracts: data, theory, and tests. Journal of Money, Credit, and Banking 29, 517–534.

Slovin, M.B., Sushka, M.E., Polonchek, J.A., 1993. The value of bank durability: borrowers as bank stakeholders. Journal of Finance 48, 289–302.

Titman, S., Wessels, R., 1988. The determinants of capital structure choice. Journal of Finance 43, 1–19.

von Thadden, E.L., 1995. Long-term contracts, short-term investment, and monitoring. Review of Economic Studies 62, 557–575.

von Thadden, E.L., 1998. Asymmetric information, bank lending, and implicit contracts: the winner's curse. Unpublished working paper. University of Lausanne (DEEP).

THE JOURNAL OF FINANCE • VOL. XLIX, NO. 1 • MARCH 1994

The Benefits of Lending Relationships: Evidence from Small Business Data

MITCHELL A. PETERSEN and RAGHURAM G. RAJAN*

ABSTRACT

This paper empirically examines how ties between a firm and its creditors affect the availability and cost of funds to the firm. We analyze data collected in a survey of small firms by the Small Business Administration. The primary benefit of building close ties with an institutional creditor is that the availability of financing increases. We find smaller effects on the price of credit. Attempts to widen the circle of relationships by borrowing from multiple lenders increases the price and reduces the availability of credit. In sum, relationships are valuable and appear to operate more through quantities rather than prices.

IN A FRICTIONLESS CAPITAL MARKET, funds will always be available to firms with positive net present value investment opportunities. In practice, managers of small firms often complain of not being able to borrow enough capital at reasonable rates. Economic theorists (for example, see Stiglitz and Weiss (1981)) suggest that market frictions such as information asymmetries and agency costs may explain why capital does not always flow to firms with profitable investment opportunities. Developing on this theme, other economists (see Leland and Pyle (1977), Campbell and Kracaw (1980), Diamond (1984), Fama (1985), Haubrich (1989), and Diamond (1991)) describe how large institutional creditors can (partially) overcome these frictions by producing information about the firm and using it in their credit decisions. If scale economies exist in information production, and information is durable and not easily transferred, these theories suggest that a firm with close ties to financial institutions should have a lower cost of capital and greater

*Graduate School of Business, University of Chicago. Petersen thanks the Center for Research on Securities Prices while Rajan thanks the Graduate School of Business at the University of Chicago for funding. We thank Andrew Alford, Bob Aliber, Douglas Diamond, William Dunkelberg, Philip Dybvig, Anne Grøn, Oliver Hart, Steven Kaplan, Anil Kashyap, Randy Kroszner, Chris Lamoureaux, Mark Mitchell, David Rudis, and Rob Vishny for valuable comments. We thank René Stulz, the editor, and an extremely thoughtful referee for many valuable suggestions. We also benefitted from the comments of participants at the NBER Summer Workshop on Corporate Finance, the Finance Workshops at the University of Chicago and Washington University, St. Louis, the Financial Markets Group at the London School of Economics, and the Western Finance Association Meetings.

availability of funds relative to a firm without such ties.[1] We term these ties relationships.

In recent years, a number of empirical studies have investigated the benefits of firm-creditor relationships. In a series of papers, Hoshi, Kashyap, and Scharfstein (1990a, 1990b, 1991) find that firms in Japan with close ties to their banks are less likely to be liquidity constrained in their investments than firms that do not have such ties. Furthermore, firms with close ties are more able to invest when they are financially distressed, suggesting again that banking relationships help overcome frictions impeding the flow of credit. For the United States, James (1987), Lummer and McConnell (1989), and James and Wier (1990) find that the existence or renewal of a banking relationship is a positive signal to the stock market. Shockley and Thakor (1992) find a similar effect for loan commitments.

Our paper differs from the ones cited above in that we use more detailed measures of the strength of firm-creditor relationships. Furthermore, we estimate the effects of relationships on both the availability and the price of credit. To the extent that we can do so accurately, we provide evidence on the precise channel or channels through which relationships benefit the firm.[2]

The data we use are from the National Survey of Small Business Finance collected by the U.S. Small Business Administration (SBA). The sample is well suited for our purposes. Only firms with fewer than 500 employees were included in the sample. The firms have a median size of book assets of $130,000 and median sales of $300,000. Since these firms are small, they are unlikely to be monitored by rating agencies or the financial press. As a result, there may be large information asymmetries between these firms and potential public investors. Furthermore, most of these firms are relatively young, with a median age of 10 years. In comparison, firms in the largest decile of New York Stock Exchange stocks have been listed for a median of at least 33 years. Since the youngest firms in our sample do not have much of a track record, a potential lender is uncertain about the competence and trustworthiness of the management, as well as the kinds of investment opportunities that could arise. If lenders remain at arm's length, management can indulge in pet projects, shift risk toward the fixed claim creditors, or otherwise misuse the borrowed funds. Some theorists have argued this is why small and young firms can rarely borrow in the public capital markets, and why we would expect firm-creditor relationships to be especially important in this sample (Diamond (1991)).

Apart from being an ideal testing ground for the theory, small firms are an important component of the national economy, producing 38 percent of gross

[1] Roosa (1951) appears to be the first to discuss the effect of bank-customer relationships in an environment with credit rationing. The recent theoretical developments discussed above have rekindled interest in the issue after a long hiatus. It should be noted that there are a few theorists who do not agree that stronger bank-creditor relations will always increase a firm's access to capital (for an example, see Blackwell and Santomero (1982)).

[2] Berger and Udell (1992), use the same data set as we do and find that a lender is less likely to demand collateral if a firm has had a long relationship with it.

national product (Dennis, Dunkelberg, and Van Hulle (1988)) and employing half of the work force (Brown, Hamilton, and Medoff (1990)). Some of these firms may be the industrial giants of the future. An important measure of the efficiency of a financial system is the extent to which such firms are nurtured and have access to the capital necessary for growth. This study is also a step toward understanding that process.

In the next section we discuss how, in theory, relationships can reduce frictions in the flow of capital from potential lenders to borrowers. This provides the basis for defining our relationship variables. Section II describes the borrowing patterns of small firms as they grow older and larger. Small firm borrowing is heavily concentrated among a few lenders, with banks being the predominant source. In Section III we examine the empirical determinants of the interest rate on the firm's most recent loan, and in Section IV the determinants of the availability of credit. This study provides evidence that relationships increase the availability and reduce the price of credit to firms. Furthermore, firms appear to reap the benefits of relationships more from increases in the quantity of finance made available by institutional lenders than through reductions in its price. Section V concludes with policy implications.

I. Theories

In most markets, prices adjust to equate demand and supply. It has been argued that the capital market is special in that the interest rate need not always adjust to clear the market. Stiglitz and Weiss (1981) show that the rate charged, to an ex ante observationally equivalent group of borrowers, determines not only the demand for capital but also the riskiness of the borrowers. A higher interest rate either draws riskier applicants (the adverse selection effect) or influences borrowers to choose riskier investments (the incentive or moral hazard effect). If an increase in the interest rate increases the average riskiness of borrowers, lenders may optimally choose to ration the quantity of loans they grant rather than raise the rate to clear the market.

As discussed earlier, adverse selection and moral hazard may have a sizeable effect when firms are young or small, which may explain why they find it hard to raise money in the public markets. However, through close and continued interaction, a firm may provide a lender with sufficient information about, and a voice in, the firm's affairs so as to lower the cost and increase the availability of credit. We term this interaction a relationship. We now examine its various dimensions.

An important dimension of a relationship is its duration. The longer a borrower has been servicing its loans, the more likely the business is viable and its owner trustworthy (Diamond (1991)). Conditional on its past experience with the borrower, the lender now expects loans to be less risky. This should reduce its expected cost of lending and increase its willingness to provide funds. It is possible that the lender could obtain sufficient information on the firm's ability to service debt-like claims by observing its past interactions with other fixed claim holders like employees or prior creditors. If so, the age of the firm rather than the length of the financial relationship should determine the lender's cost and the availability of funds. Alternatively, the information generated within a financial relationship may not be observable (or transferable) to outsiders. If so, the length of the relationship should exert an independent influence.

In addition to interaction over time, relationships can be built through interaction over multiple products. Borrowers may obtain more than just loans from a lender, especially if the lender is a bank. Firms can purchase a variety of financial services from their lender and also maintain checking and savings accounts with it. These added dimensions of a relationship can affect the firm's borrowing in two ways. First they increase the precision of the lender's information about the borrower. For example, the lender can learn about the firm's sales by monitoring the cash flowing through its checking account or by factoring the firm's accounts receivables. Second, the lender can spread any fixed costs of producing information about the firm over multiple products. Both effects reduce the lender's costs of providing loans and services, and the former effect increases the availability of funds to the firm.

We have argued above that relationships can reduce the lender's expected cost of providing capital. Whether the cost savings are passed along in the form of lower loan rates, however, depends on how competitive the capital market for small firms is. The state of competition depends, of course, on the number of potential lenders in the market and on how informed they are. If, as discussed earlier, the information generated in prior relationships can be verified by potential new lenders, they can compete on par with the current lender. If the information cannot be verified by new lenders, the current lender acquires an informational monopoly over the firm. Greenbaum, Kanatas, and Venezia (1989), Sharpe (1990), and Rajan (1992) argue that this allows the current lender to extract the rents attributable to knowing that the borrower is less risky than average. Hence, if the information generated in the relationship is private to the lender and not transferable by the borrower to others, the relationship reduces the interest rate by less than the true decline in cost. Even though these theories imply that the effect of close firm-creditor ties on the cost of funds is ambiguous, in general, the availability of funds should increase.[3]

II. Data

A. Sample Description

The data in this study are obtained from the National Survey of Small Business Finances. The survey was conducted in 1988 and 1989 under the guidance of the Board of Governors of the Federal Reserve System and the

[3] If the (ex post) monopoly distorts the firm's investment incentives excessively, availability of funds could decrease (see Rajan (1992)). If the bank can freely dispose of its monopoly power, for example with loan commitments, availability will always increase.

Table I

Distribution of Sample Firms by Industry

This table contains the distribution of firms in our sample by the one-digit SIC code.

Industry	Number of Firms	Asset Size (in 1,000s of Dollars) Min.	Mean	Median	Max.	Firm Age (in Years) Mean	Median
Mining	26	30	3,129	464	32,317	12.5	7.0
Construction	447	1	708	103	12,000	14.2	12.0
Manufacturing	408	1	2,839	452	154,087	16.4	12.0
Utilities and transportation	117	7	1,778	275	62,983	13.3	10.0
Wholesale trade	344	1	1,671	302	35,945	15.1	12.0
Retail trade	930	1	589	114	22,820	12.2	9.0
Insurance and real estate	194	1	692	153	10,671	15.7	12.0
Services	938	1	591	82	69,073	13.8	10.0

SBA. It targeted nonfinancial, nonfarm small businesses which were in operation as of December, 1987.[4] Financial data were collected only for the last fiscal year. The sample was stratified by census region (Northeast, North Central, South, and West), urban or rural location (whether the firm was located in a metropolitan statistical area (MSA)), and by employment size (less than 50 employees, 50 to 100 employees, more than 100 employees and less than 500 employees (the maximum size in the sample)). The stratification was done to insure that large and rural firms are represented in the sample. The response rate was 70 to 80 percent, depending upon the section of the questionnaire considered.

There are 3,404 firms in the sample, of which 1,875 are corporations (including S corporations) and 1,529 are partnerships or sole proprietorships. Nearly 90 percent of these firms are managed by the owner or owners. Twelve percent are owned by women and 7 percent by minorities. Small firms are concentrated in businesses that require less capital assets. Nearly 28 percent of the firms in our sample are in the service industry. These firms are the smallest when measured on the basis of the book value of assets (see Table I). Another 27 percent of the firms are in the retail trade industry. The largest firms on the basis of book assets are manufacturing firms, which comprise 12 percent of our sample.

B. *Firm Borrowing Patterns*

Before turning to the impact of relationships on the financing of small firms, we describe the pattern and sources of borrowing for firms in our sample. The corporations are significantly larger than the proprietorships or partnerships. The mean book value of assets for corporations is $1.7 million

[4]Firms involved in the agriculture, forestry, and fishing industries, finance and insurance underwriting, or real estate investment trusts were excluded from the survey.

compared to \$0.25 million for sole proprietorships and partnerships. Controlling for firm size, the corporations and noncorporations appear equally levered. The institutional debt-to-asset ratio (institutional debt excludes debt from the owners or their families) is almost identical—27 percent for corporations versus 24 percent for sole proprietorships and partnerships. These ratios conceal the large difference in the fraction of firms that have no debt. Twenty-eight percent of the corporations and 45 percent of noncorporations (sole proprietorships and partnerships) have no institutional borrowing.[5] Although more corporations have external debt financing, conditional on having institutional debt they have less leverage. The institutional debt-to-asset ratio, conditional on having institutional debt, is 43 percent for noncorporations versus 37 percent for corporations.

For firms with debt, Table II, Panel A shows the average borrowing from different sources when firms are grouped by size (book value of assets). The smallest 10 percent of firms in our sample borrow about 50 percent of their debt from banks.[6] Another 27 percent comes from the firm's owners and their families. The table shows that the fraction from personal (owner and family) sources declines to 10 percent for the largest 10 percent of firms in our sample. The fraction from banks increases to 62 percent for this group. There is no clear variation of borrowing with firm size for the other sources.

With the growing deregulation in the eighties, the distinction between banks and other financial institutions is perhaps not as clear as it once was. Therefore, we classify institutions as close if the firm obtains at least one financial service from it. Financial services include depository services—like providing checking and savings accounts—and services that provide the lender information about the firm's business—like cash management services, bankers acceptances, credit card processing, pension fund management, factoring, or sales financing. This classification attempts to capture how close the working relationship between the financial institution and the firm is. Approximately half of the institutional borrowing comes from close lenders. The fraction of institutional loans from close institutions increases from 0.45 to 0.62 as firm size increases.

Table II, Panel B describes the variation of corporate borrowing with firm age where age is defined as the number of years under current ownership.[7] The youngest firms (age less than or equal to 2 years) rely most heavily on loans from the owner and his or her family. These firms also use bank loans. In their initial years, their largest incremental source of funds is from banks, while they secularly reduce their dependence on personal funds. Eventually

[5] Firms may have unused credit lines—these would not show up in our loan volume data.

[6] We classify commercial banks, savings and loans associations, savings banks, and credit unions as Banks. Finance companies, insurance companies, brokerage or mutual fund companies, leasing companies and mortgage banks are classified as Nonbank Financial Institutions. We also have loans made by nonfinancial firms. The remaining loans consist of venture capitalist loans, loans from government agencies, and otherwise unclassified loans.

[7] We also measure age as the number of years since the firm was founded and obtained similar results.

Table II

Amount and Sources of Borrowing: By Size and Age

The first row is based on the smallest 10 percent of firms (book assets of less than \$15,000). The asset percentiles are based on the entire sample (N = 3,404). The average debt is calculated for firms with debt only. The fraction of total borrowing, from different sources is described for firms that have debt. These percentages do not sum to 100 percent since the "not otherwise classified" category is not included. The F-statistic tests the equality of the means in each column. The last column contains the fraction of debt which firms obtain from close lenders. Close lenders are institutions which also provide the firm with at least one financial service. These include checking and savings accounts, cash management services, bankers acceptances, credit card processing, pension fund management, factoring, or sales financing.

Book Value of Assets (\$1,000)	Assets Percentile	Percentage of Firms with Debt	Debt (\$1,000) Mean	Debt (\$1,000) Median	Fraction Borrowed from Each Source: Bank	Nonbank Financial Institution	Owner	Family	Other Firms	Fraction of Institutional Debt from Close Lenders
Panel A: Sources of Borrowing: By Size										
Less than 15	0–10	0.34	9	6	0.51	0.10	0.13	0.14	0.04	0.45
15–46	10–25	0.55	17	12	0.56	0.12	0.11	0.09	0.03	0.52
46–130	25–50	0.71	36	28	0.58	0.11	0.09	0.11	0.03	0.55
130–488	50–75	0.82	107	80	0.55	0.11	0.11	0.09	0.04	0.54
488–2,293	75–90	0.91	438	300	0.60	0.12	0.10	0.05	0.03	0.61
Over 2,293	90–100	0.91	2933	1585	0.62	0.14	0.07	0.03	0.04	0.62
F-statistic					2.10	0. 40	1.59	6.07	0.21	2.74
p-value					0.06	0.85	0.16	0.00	0.96	0.02
Panel B: Sources of Borrowing: By Age										
Less than 2	0–10	0.79	648	40	0.49	0.09	0.10	0.17	0.05	0.52
2–5	10–25	0.77	395	61	0.54	0.12	0.14	0.11	0.03	0.60
5–10	25–50	0.77	334	53	0.58	0.11	0.10	0.09	0.04	0.57
10–19	50–75	0.74	410	54	0.63	0.12	0.08	0.05	0.03	0.57
19–30	75–90	0.71	695	96	0.60	0.14	0.10	0.04	0.04	0.55
Over 30	90–100	0.59	912	128	052	0.15	0.10	0.06	0.04	0.50
F-statistic					5.72	1.19	2.24	13.10	0.97	1.24
p-value					0.00	0.31	0.05	0.00	0.44	0.29

firms reduce their dependence on banks too. The fraction of borrowing from banks declines from 63 percent for firms aged 10 to 19 years to 52 percent for the oldest firms in our sample (see Table II, Panel B). This seems to suggest that firms follow a "pecking order" of borrowing over time, starting with the closest sources (family) and then progressing to more arm's length sources.[8] The fraction of institutional loans from close institutions is also consistent with this observation. Except for the first group, which contains firms which are larger than average, loans from close institutions decrease as the firm gets older, from 0.60 to 0.50.[9]

C. Concentration of Borrowing

Another measure of the closeness of a borrower to its lenders is the concentration of the firm's borrowing across lenders. Table III, Panel A describes the average fraction of total firm borrowing that comes from the largest single lender when firms are grouped by size. The smallest 10 percent of firms who have a bank as their largest single lender secure, on average, 95 percent of their loans (by value) from it. By contrast, the largest 10 percent of firms obtain 76 percent of their loans from the bank. Thus, firms tend to concentrate their borrowing from one source, though this concentration decreases as firm size increases. As the table shows, such concentrated borrowing is not restricted to firms that have a bank as their largest lender. The same pattern appears no matter what the identity of the largest lender. Another way of measuring concentration is the number of sources from which a firm borrows. On average, the smallest firms tend to have just over one lender while the largest firms have about three lenders (numbers not in table).

Table III, Panel B describes the average fraction of total firm borrowing that comes from the largest single lender when firms are grouped by age. The high concentration of borrowing is still apparent in this table, but there is little variation with age. When the largest single lender is a bank, there is a slight decrease in dependency as firms age. In summary, the data show that small firm borrowing is highly concentrated. Firms diversify their sources as they become larger. It is less clear that age has any effect on diversification. Concentration of borrowing could be one measure of how close a firm is to its main lender. We will shortly describe other measures of closeness and their effect on the cost and availability of capital.

[8]The youngest 10 percent of firms in our sample borrow an amount equal to 0.32 of their book assets, while the oldest 10 percent of firms in our sample borrow only 0.15. The smallest 10 percent of firms in our sample borrow 0.22 of their book assets while the largest 10 percent of firms in our sample borrow 0.30 of their book assets. Thus, leverage decreases with age, but increases with size. A natural explanation for this is that young firms are externally financed while old firms finance via retained earnings. Larger firms may also be firms that have grown faster and have thus borrowed more.

[9]A regression shows that the fraction borrowed from close institutions is positively related to size and negatively related to the age of the firm. Both coefficients are statistically significant at the 5 percent level.

Table III

Concentration of Borrowing: By Size and Age

The results are reported by firm size, firm age, and primary source of debt. The asset percentiles are based on the entire sample and are the same as the ones used in Table II, Panel A. The first row contains information on the smallest 10 percent of the firms. The fraction of total borrowing from the largest single lender is reported by type of largest lender. The number of firms in each cell are reported in parentheses. The F-statistic tests the hypothesis that the percentage of borrowing from the largest lender is constant across the different size categories.

Panel A: Concentration of Borrowing by Size						
		Fraction of Borrowing from Largest Lender				
Book value of Asset ($1,000)	Asset Percentiles	Bank	Nonbank Financial Institution	Owner	Family	Other Firms
Less than 15	0–10	0.95 (51)	0.93 (10)	0.97 (12)	0.93 (14)	0.79 (5)
15–46	10–25	0.93 (153)	0.88 (31)	0.92 (32)	0.90 (26)	0.88 (9)
46–130	25–50	0.88 (359)	0.81 (62)	0.87 (52)	0.84 (65)	0.87 (17)
130–488	50–75	0.84 (390)	0.79 (72)	0.74 (81)	0.81 (65)	0.85 (26)
488–2,293	75–90	0.79 (296)	0.74 (49)	0.73 (38)	0.81 (23)	0.73 (15)
Over 2,293	90–100	0.76 (211)	0.72 (43)	0.75 (18)	0.74 (7)	0.71 (9)
F-statistic		19.98	2.61	6.16	1.79	1.44
p-value		0.00	0.03	0.00	0.12	0.22
Panel B: Concentration of Borrowing by Age						
Less than 2	0–10	0.86 (150)	0.80 (25)	0.85 (25)	0.89 (51)	0.76 (16)
2–5	10–25	0.85 (219)	0.77 (43)	0.82 (52)	0.83 (44)	0.86 (10)
5–10	25–50	0.85 (347)	0.76 (58)	0.81 (53)	0.83 (54)	0.84 (24)
10–19	50–75	0.85 (426)	0.80 (71)	0.80 (50)	0.78 (30)	0.88 (15)
19–30	75–90	0.82 (216)	0.84 (42)	0.73 (34)	0.77 (12)	0.70 (10)
Over 30	90–100	0.80 (106)	0.78 (28)	0.83 (19)	0.97 (9)	0.83 (6)
F-statistic		1.51	0.62	0.96	2.10	1.25
p-value		0.18	0.68	0.44	0.07	0.29

III. The Cost of Capital

A. Description of Loan Rates

In this section we examine the effect of relations on the firm's cost of debt. The data set includes the interest rate on the firm's most recent loan for 1,389 firms. The source of the loan is from institutions—a bank, a nonbank financial firm, or a nonfinancial firm—so that loans from the owner or her family are not included in this subsample. Banks are the dominant source of external capital, accounting for 82 percent of the loans in this sample. The interest rates average 11.3 percent with a standard deviation of 2.2 percent. This is 4.1 percent above the rate on a government bond of similar maturity, 2.4 percent above the prime rate at the time the loans were made, and 13 basis points below the yield on BAA corporate bonds (a basis point is one hundredth of a percentage point).

B. Determinants of the Loan Rate

Before we turn to the role of relationships, it is important that we control for the underlying cost of capital as well as loan- and firm-specific characteristics that influence the rate. In the regression results below, we use the prime rate to control for changes in the underlying cost of capital. The prime rate includes the risk-free rate and a default premium for the bank's best customers. If these small businesses are not the bank's best customers, they will pay an additional default premium. We control for aggregate variations in this premium by including the difference between the yield on corporate

Table IV

Borrowing Costs and the Role of Relationships

The dependent variable is the interest rate quoted on the firm's most recent loan. Standard errors are reported in parentheses. In addition to the variables reported, each regression also includes seven industry dummies based on the one-digit SIC codes, three regional dummies, six dummy variables for the type of assets with which the loan is collateralized, and an intercept.

Variable	(1)	(2)	(3)	(4)
Interest rate variables				
Prime rate	0.278*	0.282*	0.312*	0.278*
	(0.030)	(0.030)	(0.035)	(0.030)
Term structure spread	−0.019	−0.017	0.000	−0.027
	(0.083)	(0.083)	(0.100)	(0.083)
Default spread	0.333**	0.340**	0.183	0.325**
	(0.149)	(0.149)	(0.175)	(0.149)
Firm characteristics				
Log(book value of assets)	−0.254*	−0.264*	−0.255*	−0.259*
	(0.045)	(0.044)	(0.056)	(0.045)
Debt book assets	0.005	−0.015	−0.051	0.001
	(0.143)	(0.143)	(0.159)	(0.143)
Borrower is a corporation (0, 1)	−0.238***	−0.229***	−0.257	−0.243***
	(0.139)	(0.139)	(0.169)	(0.140)
Sales growth (1986–87)[a]			−0.585*	
			(0.301)	
Profits/interest[a]			−0.010*	
			(0.006)	
Mean 1987 gross profits/assets ratio in two-digit SIC industry				1.391**
				(0.700)
Mean σ(gross profits/assets) between 1983–87 in two-digit SIC industry				−0.771
				(0.681)
Loan characteristics				
Floating rate loan (0,1)	−0.463**	−0.448**	−0.469**	−0.473*
	(0.181)	(0.181)	(0.222)	(0.182)
Bank loan (0, 1)	0.238	0.216	0.341	0.248
	(0.225)	(0.225)	(0.270)	(0.225)
Nonfinancial firm loan firm (0, 1)	−1.125*	−1.178*	−0.513	−1.138*
	(0.360)	(0.361)	(0.430)	(0.360)

Table IV—*Continued*

Variable	(1)	(2)	(3)	(4)
Relationship Characteristics				
Length of relationship (in years)[b]	0.002	0.081	0.003	0.002
	(0.006)	(0.059)	(0.007)	(0.006)
Firm age (in years)[b]	−0.014**	−0.227*	−0.011***	−0.014**
	(0.006)	(0.078)	(0.007)	(0.006)
Information financial service (0, 1)	−0.089	−0.087	0.057	−0.087
	(0.159)	(0.158)	(0.185)	(0.159)
Noninformation financial service (0, 1)	−0.097	−0.101	−0.134	−0.104
	(0.153)	(0.153)	(0.181)	(0.153)
Deposit accounts with current lender (0, 1)	0.064	0.008	−0.041	0.061
	(0.182)	(0.186)	(0.225)	(0.182)
Number of banks from which firm borrows	0.306*	0.321*	0.303*	0.302*
	(0.085)	(0.085)	(0.096)	(0.085)
Herfindahl index for financial institutions (1, 2, or 3)	0.042	0.033	−0.024	0.033
	(0.077)	(0.077)	(0.091)	(0.077)
Number of observations	1,389	1,389	978	1,389
Adjusted R^2	0.145	0.146	0.158	0.146
Root mean squared error	2.18	2.18	2.16	2.18

[a] When profits are negative, "Profits/interest" was coded as zero. Both "Profits/interest" and "Sales growth" are truncated at their 95th percentiles (76.0 and 1.0) to limit the influence of outliers.

[b] We replace length of relationship and firm age by the natural log of one plus the length of relationship and firm age in column 2. Thus the coefficient measures the change in the interest rate due to a one percent increase in the independent variable.

* Significant at the 1 percent level.

** Significant at the 5 percent level.

*** Significant at the 10 percent level.

bonds rated BAA and the yield on ten-year government bonds.[10] We also include a term premium, defined as the yield on a government bond of the same maturity as the loan minus the Treasury bill yield, to account for interest rate differences across different loan maturities. For floating rate loans this variable is set to zero. We estimate an ordinary least squares regression of the form:

$$\begin{aligned}&\text{Interest rate on most recent loan}\\&\quad = \beta_0 + \beta_1 \text{ Economy wide interest rate variables}\\&\qquad + \beta_2 \text{ Firm characteristics} + \beta_3 \text{ Loan and lender characteristics}\\&\qquad + \beta_4 \text{ Region and industry dummies}\\&\qquad + \beta_5 \text{ Relationship characterestics} + \varepsilon. \qquad (1)\end{aligned}$$

The regression that explains the variation in the rate quoted on the most recent loan is reported in Table IV, column 1. A significant fraction of the rate

[10] We obtain the yields on government bonds from the CRSP Fama-Bliss Bond Files. We obtain the yield on BAA corporate bonds from the Citibase database.

variation is explained by economy-wide factors. The change in the loan rate due to a change in the market rate is, however, significantly less than one. A one percent increase in the prime rate raises the loan rate by 28 basis points. The relative insensitivity of the loan rate is consistent with evidence from markets for consumer borrowing (see Ausubel (1992)). Increases in the default premium also raise the firm's borrowing rate. Each percentage increase in the spread between the BAA corporate rate and the long-term government bond rate raises the average loan rate by 33 basis points.

To control for variation in the loan rate due to the characteristics of the firm we include the firm's size (book value of assets), leverage, dummies for the firm's industry (coefficients not reported), and whether the firm is incorporated. The coefficient estimates for the firm characteristics are consistent with these variables being proxies for risk. Larger firms pay lower interest rates. A firm with assets of $740,000 (the 75th percentile) can expect to pay 0.59 percentage points less than a firm with assets of only $72,000 (the 25th percentile). Being incorporated lowers the interest rate by an additional 24 basis points.

To control for variation in the loan rate due to the characteristics of the loan we include dummies for whether it is a floating rate loan, for the kind of collateral offered (coefficients not reported), and for the type of lender making the loan. We also include regional dummies, industry dummies (coefficients not reported), and a measure of the Herfindahl index of the concentration of depository institutions in the area where the firm is headquartered.

C. The Role of Relationships

Based on the discussion in Section II, we expect relationships to lower the lender's cost of lending to small firms. We estimate the effect of relationships on the interest rate charged. Implicit, therefore, in our analysis is the assumption that reductions in the lender's cost are passed on to the borrower in a lower rate. The first dimension of a relationship that we include is the length of the relationship between the borrower and its current lender. This should be a proxy for the private information the institution has about the firm. Firms who have been doing business with their lender for a short time should pay a higher rate. Of course, we must distinguish this effect from the fact that younger firms pay higher rates on their loans (Dennis, Dunkelberg, and Van Hulle (1988)). The length of the relationship and the age of the firm are correlated but not as highly as expected ($\rho = 0.41$). When both variables are included in the regression, we find little independent importance for the length of the relationship (see Table IV, column 1). The coefficient is positive, but its magnitude is statistically zero ($\beta = 0.002$, $t = 0.3$). Older firms, however, are charged statistically smaller interest rates; an additional year lowers the interest rate by 1.4 basis points or 0.014 percentage points ($t = -2.3$).

The firm's reputation may not increase linearly with the age of the firm. The effect of an additional year of existence should decline with the age of the firm. To test for a possible nonlinear relation, we first estimate a separate slope for the firm age variable when firm age is less than 10 years (the median age). The coefficient is slightly larger, but the larger standard error means the coefficient is not statistically different from zero. We next replace the firm age and the relationship age by the log of one plus the age. This allows the marginal impact of age to decline. The estimates are reported in Table IV, column 2. The coefficient on the length of the relationship is again not significantly different from zero, while the coefficient on firm age indicates a declining impact of age. An additional year reduces the interest rate by 16 basis points if the firm has just been founded or acquired (age = 0), but only 2 basis points if the firm is 10 years old. To see the economic importance of this coefficient, a one standard deviation increase in the log of one plus the age of the firm reduces the interest rate it is charged by 0.19 percentage points. Later, we will compare the economic impact of relationship proxies on the interest rate with their impact on the availability of credit. An admittedly crude comparison is to calibrate these effects against a common standard, i.e., the effect of an increase in firm size. A one standard deviation increase in the size of the firm reduces the interest rate by 0.47 percentage points. Thus the effect of firm age on the interest rate is only 40 percent as large as the effect of firm size on the interest rate.[11]

The R^2 in columns 1 and 2 is almost identical, meaning that the data do not distinguish between a linear specification and a log linear specification. We also use the alternative definition of firm age as the number of years since the firm was founded rather than the number of years under current ownership. The coefficient on firm age drops by two thirds. The owner's reputation is apparently more important than that of the business.

The second measure we examine is the nonborrowing side of the firm's relationship with its current lender. In addition to borrowing, the firm may have checking or savings deposits with its current lender. Sixty-four percent of our sample does. The firm may also purchase financial services from the firm. As discussed earlier, these nonloan services can be used by the lender to monitor the firm. If these sources of information reduce monitoring costs or improve the accuracy of the lender's information, they should reduce the expected cost of such loans. We have already listed the financial services that might provide information to the lender (see Section II.*B* for a list of these services). In addition, the lender may perform services that arguably do not give it information—for example, providing change and night depository services. We code dummy variables for whether the firm had checking or savings deposits with the current lender, whether it purchased other informationally intensive financial services from it, and whether it purchased noninformational services.

[11] The figure of 40 percent is obtained by dividing the magnitude of the effect of a one standard deviation change in the log of one plus firm age on the interest rate (0.19 percent) by the effect of a one standard deviation change in firm size on the interest rate (0.47 percent).

That a firm obtains financial services from the current lender has no significant effect on the interest rate in our sample (see Table IV, column 1). Lenders who provide their customers with informationally intensive services charge a lower rate on their loans; however, the magnitude of this reduction is tiny (9 basis points). In addition, all three coefficients are statistically indistinguishable from zero.

Our third measure of the strength of the relationship is how concentrated the firm's borrowing is. From the results in Section II, it is clear that the firms in our sample borrow a significant fraction of their debt from a single institution. Even the largest firms in our sample borrow three quarters of their debt from a single institution (see Table III, Panel A). Firms may concentrate their borrowing with a lender to reduce overall monitoring costs, improve the lender's control, and cement their relationship. In these cases, concentrated borrowing should be associated with lower cost credit. Alternatively, firms may borrow from a single lender because it is their only source of credit. If so, then concentrated borrowing should be associated with more expensive credit.

We use the number of banks from which the firm borrows as a measure of borrowing concentration.[12] The firms in our sample borrow from no more than six banks, and the median firm borrows from only one bank. Eighteen percent of the firms borrow from more than one bank. We find that the rate paid by a firm increases by a significant 31 basis points when a firm increases the number of banks from which it borrows by one (Table IV, column 1). If we use the calibration method discussed earlier, the effect of the number of banks on the interest rate is about 53 percent of the effect of size.

As an alternative measure of concentration, we include the number of nonbank institutions from which the firm borrows. Increasing the number of nonbank institutions from which the firm borrows has no effect on the firm's borrowing rate. It is perhaps more plausible to think that ties between a firm and a bank are more indicative of a close relationship than ties between a firm and a nonbank. If so, this evidence suggests that the rate increases with a multiplicity of relationships rather than a multiplicity of creditors. In summary, a single banking relationship lowers borrowing costs, while multiple banking relationships are costly.

An alternative interpretation of the above result is that the number of banks is really a proxy for the firm's quality. Lower quality firms, unable to borrow additional money from their first bank, must approach other banks for additional capital. If so, the unwillingness of the original bank to extend the firm additional credit may be a signal of the firm's riskiness or quality, and the firm can obtain credit at a second bank only at a higher rate. In discussions with small business bankers, we were told there are several reasons besides quality why a firm may have multiple banks. Some banks specialize in the type of loans they make. Thus, a firm whose management wants to borrow with a line of credit against its accounts receivable may have

[12] We also consider the fraction of the firm's debt that is borrowed from its current lender. The results are qualitatively identical.

to approach a different bank than the one that made it a mortgage loan. Some firms borrow from multiple banks, so they can play the banks off against each other. Finally, some owners like the prestige of multiple banking relationships. To test the quality hypothesis, we divide the sample into those firms that have one bank and those that have more than one bank. We then search for differences between the two samples.

The firms with multiple banks are over twice as large as those with only one bank. As firms grow, they expand the number of banks from which they borrow. But these are not necessarily firms which are in the process of expanding (over) aggressively. The firms with multiple banks have lower sales growth (16 percent versus 35 percent).[13] They also have lower interest coverage (median profits/interest of 2.2 versus 4.3). These numbers suggest that the number of banks may be a proxy for lower quality firms. To test this hypothesis we include interest coverage and sales growth as additional explanatory variables in the interest rate regression (see Table IV, column 3). Both variables help predict the interest rate, and both are marginally statistically significant.[14] But the coefficient for the number of banks is only marginally lower than that in column 1. This suggests that the number of banks is not strictly a proxy for quality.

Finally, it is possible that since the data come from a survey of small businesses, many of which may not be audited, the profit figure is uninformative. While we do not have access to the names of the firms and cannot obtain more data on them, we know the two-digit Standard Industrial Classification (SIC) industry code for each firm. From COMPUSTAT, we extract the average gross-profits-to-asset ratio in 1987 for each firm's industry. We also calculate the standard deviation of the gross-profits-to-assets ratio between 1983 and 1987 for each COMPUSTAT-listed firm and obtain the average for the two-digit industry.[15] The first is a measure of profitability, and credit quality should increase with it. The second is a measure of risk, and credit quality should decrease with this variable. We report the results in column 4 of Table IV. The coefficients have the opposite sign to that expected. The interest rate is increasing in the average profitability and declining in the variability of profitability. Only the first coefficient is significantly different from zero.[16]

[13] For most variables the survey includes financial data only for 1987. It does, however, include sales figures for both 1986 and 1987. We use these numbers to calculate the firm's sales growth.

[14] Interest rate coverage will depend in part on the interest rate of the current loan. This endogeneity will bias the coefficient downward. Thus our estimated coefficient is probably more negative than the true coefficient.

[15] We only consider COMPUSTAT firms with book value of assets in 1987 below $150 million. We consider lead and lagged average profits, but these do not enter significantly.

[16] We examine this further by dropping loans where the interest rate was below the government bond rate. Presumably, these loans are made as part of a broader set of transactions and may not represent the true (relationship-adjusted) cost. The coefficients on average industry profits and standard deviation of industry profits reverse and have the expected sign, suggesting that loans to some poor quality firms—with low industry profits and high industry standard deviation of profits—are made at rates below the risk-free rate. Petersen and Rajan (1993) explore this issue in greater detail.

Not all of our proxies for the strength of firm-lender relationships are correlated with cheaper credit. That these variables do not all have a significant effect on the observed interest rate is consistent with three different theoretical explanations and an econometric one. The simplest one is that relationships do not matter much because all information is public or, at least, easily verifiable. If any potential lender can evaluate a loan's risk as accurately (and at the same cost) as the relationship lender there is no value to a specific relationship. A second possibility is that relationships do indeed have value, but rationed firms prefer greater availability of funds to a reduction in price. A third possibility is that the lender is not compelled by market forces to pass on the benefits via a lower interest rate. If the relationship confers a monopoly on the lender, this is what we would expect. The econometric explanation is that our measures may not capture the existence or strength of relationships.

IV. The Availability of Credit

A. How to Measure the Availability of Credit

We now estimate the effect of relationships on the availability of credit. If our proxies for relationships predict the availability of credit, then the econometric problem discussed above does not explain our interest rate regression. Furthermore, we may be able to distinguish among the theoretical explanations. Unfortunately, it is difficult to measure credit availability directly. The firm's debt ratio will underestimate the credit available to the firm—firms may have low debt ratios because the firm is liquidity constrained (a supply constraint) or because they have little need for external capital (a demand constraint).

The firm's debt ratio is simultaneously determined by the firm's demand for credit and the supply of credit from institutions. Thus regressions that use the firm's debt ratio as the dependent variable will suffer from a simultaneous equations bias. Changes in the debt ratio can be due to changes in demand for credit (the supply curve is observed) or by changes in supply of credit (the demand curve is observed). This statistical problem is apparent when we regress a firm's debt-to-asset ratio on characteristics of the firm. The results are reported in Table V. The dependent variable is total debt divided by assets. Credit availability should be greater for higher quality firms. Consistent with this intuition, large firms and firms in industries with high average earnings and low earnings volatility tend to have a high debt-to-assets ratio. However older firms and more profitable firms—which should be higher quality—have lower, not higher, debt ratios. The problem is we cannot tell whether older firms are rationed by creditors (a supply effect) or whether they have a lower demand for external credit. Since the coefficient estimates from this regression are not unbiased, we propose an alternative measure of the credit available to the firm.

Table V

Determinants of the Firm's Debt Ratio

The dependent variable is total debt divided by the book value of assets. It has been multiplied by 100. Since the debt ratio is censored at zero we estimate the coefficients using a one-sided tobit model. Asymptotic standard errors are reported in parentheses. The regression also includes seven industry dummies based on the one-digit SIC codes.

Variable	
Log(book value of assets)	4.40*
	(0.56)
Profits/assets (%)	−0.99*
	(0.38)
Borrower is a corporation (0, 1)	9.32*
	(1.99)
Firm age (in years)	−0.80*
	(0.09)
Length of longest relationship	−0.19**
(in years)	(0.08)
Herfindahl index for bank deposits	3.57*
	(1.33)
Mean 1987 gross profits/assets ratio	19.34**
in two-digit SIC industry	(8.92)
Mean σ(Gross profits/assets) between 1983–87	−11.94
in two-digit SIC industry	(9.48)
Number of observations	3,233
χ^2	347.1
(p-value)	(0.000)

* Significant at the 1 percent level.
** Significant at the 5 percent level.

If institutions limit the credit extended to a firm, the firm will borrow from more expensive sources, so long as the returns from its investments exceed the cost of funds from those sources. Firms with unlimited access to institutional credit will never turn to the more expensive source. Therefore, with certain caveats discussed below, the amount borrowed from more expensive sources should measure the degree to which firms are supply constrained by institutions. More specifically, let the firm's rate of return on the marginal dollar invested be given by curve *JKE* in Figure 1. The firm should invest until the rate of return from the marginal dollar of investment equals the opportunity cost of capital. The firm has three sources of capital: internally generated cash flow (*OB*), borrowing from institutions (*BC*), and borrowing from an alternative source (*CD*).

The firm will exhaust its cheapest source, internal cash, before approaching the financial institutions. If institutions do not ration credit, the firm will invest to the point where the (increasing) marginal cost of borrowing from institutions (represented by curve *GN*) intersects the curve *JKE*. The firm will invest *OM*. If, however, institutions ration the amount of credit they offer the firm, say to amount *BC*, the firm only invests *OD*. Holding all else equal,

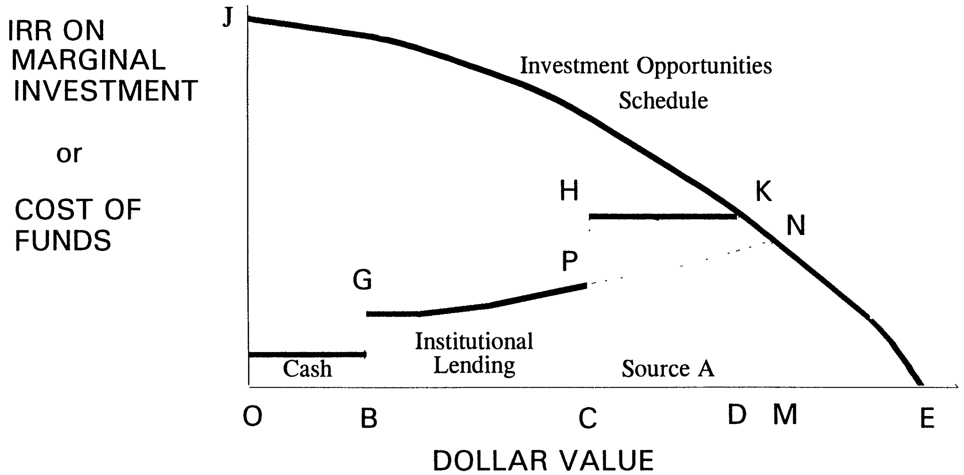


Figure 1. Sources and uses of funds. The solid curve *JKE* represents the Internal Rate of Return (*IRR*) on the marginal investment. *GP* is the marginal cost of institutional credit offered, while *HK* is the marginal cost of borrowing from *Source A*. *OB* is the amount of cash the firm has, *BC* is the amount borrowed from institutional lenders, and *CD* is the amount borrowed from *Source A*. *OD* is the total amount invested.

the amount *CD* that the firm borrows from the alternative source is then an inverse measure of the quantity of credit available from institutions. For *CD* to be an appropriate measure of institutional credit rationing, the marginal cost of borrowing from the alternative source must exceed the marginal cost of available institutional credit. If this is not true, the amount *CD* will be a function of the price financial institutions charge, as opposed to the volume of credit they are willing to offer. Also, the cost of borrowing from the alternative source should be relatively similar for firms within an identifiable class. Otherwise the amount *CD* will be a function of the specific firm's cost of borrowing from the alternative source.

What could the expensive source of credit be? Most of the firms in our sample are offered trade credit—short-term financing, which some suppliers provide with their goods and services.[17] We have data on the percentage of a firm's purchases that are made on credit, the percentage of this credit that is accompanied by discounts for early payment, the percentage of the discounts that are taken, and the percentage of trade credit that is paid late. In general, discounts for early payment and the penalties for late payment are substantial. They are meant to encourage the firm to pay on time. For

[17]Clearly, the trade creditor realizes a margin on the goods sold which is why she may be prepared to offer credit tied to the purchase of the goods even when others ration credit. Mian and Smith (1992) offer a variety of other reasons why trade creditors may do this. The manufacturer may find the collateral (merchandise) more valuable and easier to sell after repossession. Also, she may have a cost advantage in credit evaluation. Trade credit may be a method of price discrimination. Finally, there may be tax advantages if the financing qualifies as an installment loan. None of this suggests that trade credit is a cheap source of finance.

example, some firms in the retail business face the terms 10–2–30 (Smith (1987)). This is a discount of 2 percent if paid within 10 days (the "discount" date) and the full amount if paid in 30 days (the "due" date). Foregoing the 2 percent discount is equivalent to borrowing at an annual rate of 44.6 percent.[18]

Clearly, the annualized rate is not that high if firms are allowed to stretch repayments beyond the due date. Since the stated terms in an industry may differ from actual industry practice, we use our data to construct empirical measures of the actual stretch that firms face. Based on each firm's stock of accounts payable, we construct the days payable outstanding (DPO) for each firm, which is defined as 365 times the firm's accounts payable over its cost of goods sold. We report the DPO by industry in Table VI.[19] To estimate the potential stretch available to trade credit borrowers, we calculate the difference in the DPO between firms that regularly take the early payment discounts and those that do not. For each industry, we determine the median DPO for firms that take less than 10 percent of the discounts they are offered and the median DPO for firms that take more than 90 percent of the discounts offered. The difference between these two numbers is reported as the "Discount Stretch" in Table VI, and it is an estimate of how long firms that do not take discounts stretch their credit. For the retail industry it is 8.9 days.[20] Based on the standard terms, firms that do not take the discount are paying an additional 2 percent for 8.9 days of credit, which translates to an annualized interest rate of 129 percent.

A second way in which the firm can extend its trade credit financing is by paying late, i.e., after the due date. Clearly, the firm will incur both reputational and pecuniary penalties for paying late. For example, among gasoline wholesalers margins are so thin that a firm paying late may be forced to pay cash for future purchases and may be cut off from future supplies.[21] For each industry, we estimate the "Late Payment Stretch" as the difference between the median DPO for firms that repay more than 50 percent of their trade credit late and the median DPO for firms that repay less than 10 percent of their trade credit late. We find it to be 16.9 for the retail industry. Thus if the firm does not take the discount by paying on the tenth day and stretches the

[18] By taking the early payment discount, the firm is borrowing at 2/98 or 2.04 percent per 20-day period. Since there are 365/20 or 18.25 such periods in a year, this is equivalent to an annual rate of 44.6 percent ($[1 + 2/98]^{(365/20)} - 1$).

[19] We used two classifications for industry—the two-digit SIC code and the one-digit SIC code. We report only the broader classification in Table VI but use the two-digit SIC code in the estimates reported in Tables VII, VIII, and IX.

[20] Why is this number so low compared to the 20 days that should be the case if the discount terms are 2–10–30? A possible reason is that the discount date is not strictly enforced while the due date is, so that firms get discounts even if they pay after the tenth day (see Dun and Bradstreet (1970)). Another possible reason is that firms stretch entirely on the portion offered with discounts and not on any of the trade credit offered with net terms. If this is true (and we have no reason to believe that the firm should not stretch trade credit offered on net terms also), the stretch goes up to $8.9/0.3 = 30$ days. This is an implicit interest rate of 27.9 percent, which is still higher than the highest interest rate on institutional loans in our sample (24.5 percent).

[21] Authors discussions with Mr. Chuck Patton, Credit Department, Amoco Oil Company.

Table VI

Days Payable Outstanding and Stretch by Industry

Days Payable Outstanding = 365 * Firm's Accounts Payable/Costs of Goods Sold. "Discount Stretch" is defined as the difference within the one-digit SIC industry between the median days payable outstanding for firms that take fewer than 10 percent of their early payment discounts and the median DPO for firms that take more than 90 percent of their early payment discounts. "Late Payment Stretch" is defined as the difference within the industry between the median days payable outstanding between firms that pay more than 50 percent of trade credit late and firms that pay less than 10 percent of their trade credit late. Due to the small number of observations in the utilities and transportation industry, we are not able to calculate the Discount Stretch or the Late Payment Stretch.

Industry	Number of Firms	Days Payable Outstanding			Fraction of Firms Taking $\geq$ 90% of Early Payments Discounts	Discount Stretch (in Days)	Fraction of Firms Paying $\geq$ 50% of Trade Credit Late	Late Payment Stretch (in Days)
		Median	Mean	Std Dev				
Mining	20	30.6	43.7	50.8	21.4	62.0	23.1	60.4
Construction	355	15.9	37.5	96.6	62.5	22.7	22.5	21.0
Manufacturing	356	27.3	43.8	83.3	42.3	14.9	24.6	18.4
Utilities and transportation	5	9.4	18.5	20.6	66.6	—	0.0	—
Wholesale trade	296	19.7	37.2	84.6	60.4	19.9	19.8	13.6
Retail trade	825	10.8	25.2	60.9	57.6	8.9	12.8	16.8
Insurance and real estate	69	0.0	207.2	1580	84.0	1.2	15.0	96.7
Services	60	4.7	16.5	26.3	56.7	16.5	12.0	0.0

payment out for 36.9 days (20 days plus the late payment stretch of 16.9), the implicit annual interest rate is 22.1 percent.[22] This is an underestimate of the true borrowing rate because it overstates the actual discount stretch that we estimate for the retail industry (8.9 days). It also underestimates the true borrowing cost because it ignores the reputational and pecuniary costs that missing the due date will impose on the firm. Despite these omissions, this interest rate is higher than 99.8 percent of the loans in our sample.

As a final way to document the true costs of financing a firm through delayed payment of its trade credit obligations, we calculate the percent of firms in an industry that take most of the early payment discounts (more than 90 percent) and the percent of firms that pay a significant fraction of their trade credits late (more than 50 percent). These numbers are reported in Table VI. That 58 percent of firms in the retail industry avail themselves of 90 percent or more of the discounts they are offered suggests that discounts are sizeable, and that firms that forego them are not getting cheap financing. Similarly, only 13 percent of firms pay more than 50 percent of their trade credits late, suggesting that the penalties for late payment are substantial. Furthermore, our conversations and previous work (Dun and Bradstreet (1970), Elliehausen and Wolken (1992)) indicate that discount terms are not specific to a firm, but are common practice throughout an industry. As these discounts and penalties are substantial and are industry specific and not firm specific, the fraction of trade discounts not taken or the fraction of trade credit paid late are good proxies for the amount CD.

B. Trade Credit Data

In Table VII we present summary statistics for the data on trade credit. Larger (Table VII, Panel A) and older (Table VII, Panel B) firms make more of their purchases on credit, suggesting that the decision to offer credit seems to be firm specific. The percentage of credit offered with discounts for prompt payment, however, is invariant to firm characteristics like size and age. We test whether this percentage varies across age or size categories in Table VII. We cannot reject the hypothesis of a constant mean in either case ($p = 0.93$ for size and 0.63 for age). We also regress the percentage of discounts offered on several firm characteristics and 12 industry dummies. Only the industry dummies are statistically significant. It appears that once the decision to offer credit is made, discounts for early payment automatically follow if it is the supplier's policy. This evidence also seems to indicate that the size of the

[22] Neither measure of stretch is completely accurate. The discount stretch has the problems discussed in footnote 20, while the late payment stretch overestimates the stretch from the due date because it does not take into account the possibility that early payers may take substantially more of their discounts. Yet another measure of the stretch could be the difference in medians between those taking 90 percent of their discounts and those paying more than 50 percent late. In the retail industry, this is 19.4 days, which translates to a 46.2 percent annual rate.

Table VII

Trade Credit Usage of Firms: By Size and Age

Firms are grouped by size in Panel A and by age in Panel B. The number of firms in each cell are reported in parentheses. The F-statistic tests the hypothesis that the mean percentage is constant across the different size and age categories.

Panel A: Trade Credit Usage of Firms: By Size

Book Value of Assets ($1,000)	Asset Percentiles	Percentage of Purchases Made on Credit	Percentage of Trade Credit with Early Payment Discounts	Percentage of Offered Discounts Taken by Firm	Trade Credit Paid Late in Percent	Median Stretch for Firms (Measured from Last Day for Discounts)[a]	Median Stretch for Firms (Measured from Due Date)[b]
Less than 15	0–10	71.8 (203)	33.2 (204)	74.2 (122)	18.0 (100)	−13.03 (108)	−17.91 (102)
15–46	10–25	73.3 (366)	31.4 (364)	64.4 (245)	21.2 (194)	−7.32 (245)	−10.71 (239)
46–130	25–50	74.0 (639)	32.7 (632)	63.8 (457)	19.8 (344)	−2.60 (445)	−5.86 (439)
130–488	50–75	81.2 (694)	32.8 (686)	64.9 (539)	21.7 (379)	1.44 (473)	−0.81 (464)
488–2293	75–90	84.8 (446)	31.5 (433)	68.3 (375)	21.4 (260)	6.59 (305)	3.75 (308)
Over 2293	90–100	90.2 (298)	34.1 (289)	70.5 (278)	21.0 (172)	5.27 (214)	2.85 (210)
F-statistic		25.65	0.27	2.17	0.56		
p-value		0.00	0.93	0.05	0.73		

Panel B: Trade Credit Usage of Firms: By Age

Firms Age (Years)	Age Percentile						
Less than 2	0–10	75.4 (299)	33.6 (295)	58.9 (224)	25.1 (155)	−2.20 (208)	−5.86 (200)
2–5	10–25	74.3 (414)	32.2 (400)	57.1 (295)	23.7 (243)	0.00 (282)	−3.16 (276)
5–10	25–50	79.2 (623)	30.4 (615)	61.5 (457)	22.1 (357)	−0.02 (400)	−2.97 (398)
10–19	50–75	80.6 (689)	33.3 (682)	68.9 (539)	18.4 (392)	0.00 (470)	−2.71 (473)
19–30	75–90	82.7 (377)	33.4 (376)	74.5 (308)	18.6 (210)	2.89 (260)	−1.22 (252)
Over 30	90–100	83.5 (244)	34.1 (240)	82.4 (193)	15.8 (92)	0.11 (170)	−0.72 (163)
F-statistic		6.47	0.69	14.89	4.09		
p-value		0.00	0.63	0.00	0.00		

[a] For each two-digit SIC industry, the median DPO is obtained for firms availing of more than 90 percent of their discounts. This is subtracted from the DPO for the firm to obtain the stretch as measured from the last day for discounts.

[b] For each two-digit SIC industry, the median DPO is obtained for firms paying less than 10 percent of credit late. This is subtracted from the DPO for the firm to obtain the stretch as measured from the due date.

discounts offered for early payment are unlikely to be tailored to the specific customer.

The two variables of interest are the percentage of trade credit that is paid after the due date (which we call late payments) and the percentage of discounts for early payment that are taken (which we call discounts taken). Both variables are taken from the survey. Each is a proxy for the amount borrowed from the alternative source. A firm that makes more late payments or takes fewer cash discounts uses a greater amount of trade credit as a source of financing. As seen in Table VII, these two variables do not seem to depend strongly on firm size, but do depend on age. Late payments decrease from 25.1 percent for the youngest firms to 15.8 percent for the oldest firms. Discounts taken increase from 58.9 percent for the lowest age category to 82.4 percent for the oldest firms.

C. Test of the Effect of Relationships on Credit Availability

To determine if relationships increase the availability of credit, we regress late payments and discounts taken against measures of the firm's investment opportunities, its cash flow, its debt, and various measures of relationships. The regression is of the form:

$$\begin{aligned}
&\text{Trade credits paid late} \\
&\quad = \beta_0 + \beta_1 \text{ Measures of investment opportunities} \\
&\qquad + \beta_2 \text{ Industry dummies} + \beta_3 \text{ Measures of cash flow} \\
&\qquad + \beta_4 \text{ Measures of relationships} + \varepsilon. \qquad (2)
\end{aligned}$$

We include three measures of the firm's investment opportunities. Younger firms may have different investment opportunities than older firms. This may account for the pattern in Table VII, Panel B. Thus, firm age is one measure of investment opportunities. As discussed earlier, it is also a measure of the publicly available component of information. Investment opportunities could also depend on the firm's size (the book value of its assets). Finally, investment opportunities depend on the industry the firm is in, and thus industry dummies are included as explanatory variables. This will also control for differences in the price of trade credit financing across industries.

The firm's internal cash flow (normalized by book assets) is accounted for by including income after interest. While we do not have figures for depreciation, it ought to be a function of the firm's book assets which is already in the regression. We also include the ratio of outstanding institutional debt (i.e., total loans less family and owner loans) to book assets. This is a measure of the debt capacity the firm has already exhausted.[23] Finally we include a dummy for whether the firm is a corporation or not, because credit rationing

[23]An argument could be made for leaving debt out since if we perfectly control for investment opportunities, the level of trade credit used is an exact measure of the level of debt available. Leaving debt out of the regressions has no qualitative impact on the results.

should be greater for firms with limited liability. An owner-managed firm has a greater incentive to take on risky projects if it has limited liability.

We come now to the relationship variables. The first is the length of the longest relationship the firm has had with a financial institution. Second, we include a measure of how informed the firm's lenders are—this variable is the fraction of borrowing that comes from institutions that provide at least one significant financial service to the firm.[24] Third, we include a measure of how concentrated the firm's borrowing is—the number of institutions that account for more than 10 percent of the firm's borrowing. Finally, we include the Herfindahl index for financial institutions in the immediate area of the firm.[25] We cannot estimate our model with ordinary least squares since both dependent variables are expressed as percentages and are consequently censored at 0 or 100. The coefficients in a least squares estimate would be biased toward zero. We therefore estimate a tobit regression with two-sided limits.

D. Findings

The regressions explaining late payments are reported in Table VIII. The investment and cash flow variables have the predicted sign. Older and larger firms do not make as many late payments. More profitable firms do not make as many late payments, though this effect is not statistically significant.[26] Firms that have taken on more debt are more likely to pay late. Finally, corporations make more late payments. We now examine the relationship variables.

The length of the longest relationship with a financial institution is both economically and statistically significant regardless of whether we use a linear specification for firm age and relationship length (Table VIII, column 1) or a log specification (Table VIII, column 3). It is instructive to compare the economic magnitudes of the age and relationship coefficients estimated here with those estimated in the rate regression. A one standard deviation increase in the log of one plus the firm age reduces the percentage of trade credits paid late by 1.35. A one standard deviation increase in the log of one plus the length of the relationship reduces the percentage of trade credits paid late by 2.05. A one standard deviation increase in size reduces the percentage of trade credits paid late by 1.48. Following our crude method of calibration (see Section III.*C*), firm age has about 90 percent of the impact that firm size has on the availability of credit while it has only 40 percent of the impact that firm size has on the price of credit. More interesting,

[24] These are either deposit accounts or the informational financial services defined above.

[25] The survey does not report the actual Herfindahl index. We know whether the Herfindahl index is less than 0.10, between 0.10 and 0.18, or greater than 0.18. Our variable is therefore coded as 1, 2, or 3.

[26] Profits could proxy for a firm's cash flow which should reduce the amount paid late, but it could also proxy for the profitability of a firm's investment opportunities which would increase the amount paid late. The predicted effect is thus ambiguous.

Table VIII

Credit Availability and the Role of Relationships Dependent Variable: Trade Credits Repaid Late (%)

The dependent variable is the percentage of trade credits that were paid after the due date (paid late). The coefficient estimates are from a tobit regression with two-sided censoring. The dependent variable is censored at 0.0 and 1.0. Asymptotic standard errors are in parentheses. Each regression also includes seven industry dummies based on the one-digit SIC codes.

Independent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Log(book value of assets)	−0.662	−0.542	−0.753	−0.678	−0.533	−1.580*
	(0.490)	(0.495)	(0.486)	(0.513)	(0.489)	(0.566)
Profits/book assets	−0.679	−0.785	−0.663	−0.643	−0.744	0.372
	(0.753)	(0.754)	(0.753)	(0.782)	(0.750)	(0.916)
Debt from institutions/ book assets	4.853**	4.902**	4.888**	4.463**	5.775*	5.932*
	(2.040)	(2.038)	(2.045)	(2.108)	(2.045)	(2.310)
Firm is a corporation (0, 1)	2.819***	2.857***	2.827***	3.248***	2.597	5.267*
	(1.635)	(1.633)	(1.638)	(1.726)	(1.632)	(1.936)
Firm age (in years)[a]	−0.142***	−0.150***	−1.531	−0.181**	−0.140***	−0.164***
	(0.081)	(0.081)	(1.080)	(0.084)	(0.081)	(0.093)
Length of longest relationship (in years)[a]	−0.155**	−0.152**	−2.299**	−0.165**	−0.150**	−0.177**
	(0.069)	(0.069)	(1.089)	(0.070)	(0.068)	(0.078)
Debt from financial service provider (%)	−5.576*		−5.385*	−5.555*	−5.713*	−6.976*
	(1.672)		(1.678)	(1.748)	(1.663)	(1.907)
Debt from financial service provider (only one service) (%)		−7.630*				
		(2.085)				
Debt from financial service provider (multiple services) (%)		−3.895**				
		(1.958)				
Number of institutions from which firm borrows	1.926**	1.872**	1.991**	2.286**	1.840**	1.546
	(0.900)	(0.900)	(0.901)	(0.946)	(0.897)	(1.043)
Herfindahl index for financial institutions (1, 2, or 3)	−2.253*	−2.300*	−2.274*	−2.065**	−1.901**	−2.659*
	(0.866)	(0.865)	(0.866)	(0.902)	(0.866)	(0.996)
Sales growth (1986–1987)				−1.445		
				(1.086)		
Mean 1987 gross profits/assets ratio in two-digit SIC industry					−28.245*	
					(8.643)	
Mean σ(gross profits/assets) between 1983–87 in two-digit SIC industry					22.677*	
					(7.851)	
Late payment stretch (in days)[b]						0.012
						(0.019)
Fraction of firms paying $\geq$ 50% of trade credit late						84.997*
						(5.758)
Number of observations	1,119	1,119	1,119	1,019	1,119	857
χ^2	83.5	86.2	80.7	86.3	96.6	127.1
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

[a] We replace length of relationship and firm age by the natural log of one plus the length of relationship and firm age in column 3. Thus the coefficient measures the change in the interest rate due to a one percent increase in the firm's age or the length of its longest relationship.

[b] For each two-digit SIC industry, the median DPO is obtained for firms paying less than 10 percent of credit late. This is subtracted from the DPO for firms paying more than 50 percent of credit late to obtain the late payment stretch.

* Significant at the 1 percent level.

** Significant at the 5 percent level.

*** Significant at the 10 percent level.

relationship length has about 138 percent of the impact that firm size has on the availability of credit while it has no impact on the price of credit.

Firms are less likely to pay late when their lenders are more informed. The coefficient on the fraction of debt from institutions that provide financial services is -5.6 ($t = 3.3$). If the provision of services is a good measure of the closeness of the lending relationship, then lenders who provide more services are closer and should increase availability even more. This is indeed the case (Table VIII, column 2). A firm can reduce late payments by increasing the fraction it borrows from an institution providing a single service ($\beta = -3.9$), but increasing the fraction borrowed from an institution providing two or more services has almost twice the effect ($\beta = -7.6$). Providing more information to lenders has little effect on the price of credit (see Section III), but it significantly increases its availability.

In Section III we found that concentrated borrowing is correlated with cheaper credit. It is also correlated with greater availability of credit. An increase of one in the number of institutions from which the firm borrows increases late payments by almost two percentage points (Table VIII, column 1). When banks and nonbanks are considered separately, the effect of an increase in the number of banks is statistically and economically more important than an increase in the number of nonbank institutions. The coefficients are 2.5 versus 1.8, although we do not report this regression in the table. Finally, following our calibration, the number of banks has 142 percent of the impact on the availability of credit that size has. Recall that in Section III, we found the number of banks to have only 53 percent of the impact that size has on the price of credit.

Interestingly, credit availability for firms in more geographically concentrated banking markets is significantly higher. A firm in the most concentrated area reduces late payments by 4.6 percentage points when compared to a firm in the most competitive area. By comparison, concentration of the local financial market has only a small and statistically insignificant effect on the price of credit (see Table IV). Petersen and Rajan (1993) explore this issue in greater detail. They argue that a possible reason banks help out small firms is because of the possibility that such firms generate significant future business when they grow. In return for a stake in the firm's future, the bank lends even when no one else will. Unfortunately, the firm cannot explicitly commit to giving the bank a stake, because banks in the United States are statutorily prohibited from holding equity in firms. This implies that the bank has to rely on an implicit promise that it will receive the firm's future business. In a concentrated market, such a promise is more credible because the firm has few options (until it grows large enough to approach arm's length markets). They find evidence consistent with such an explanation.

According to our hypothesis, firms could finance themselves with greater amounts of expensive trade credit, not just when institutions restrict their access to credit but also when they have better investment opportunities. A potential problem with our results is that we may not be measuring investment opportunities correctly. If firms with good investment opportunities are relatively young, have short relationships, and use multiple lenders to fund their investments, we would find that all three variables are correlated with our measures of usage of trade credit. Under the assumption that high-growth firms have above average investment opportunities, sales growth is a proxy for investment opportunities. If our relationship variables are better proxies for investment opportunities than for relationships, the inclusion of sales growth in the regression should reduce the magnitude of the coefficients dramatically. We report the coefficients in the fourth column of Table VIII. Two of the three relationship coefficients *increase* in magnitude. The coefficient on the fraction of debt from institutions that provide financial services decreases slightly. We find similar results when we use book assets to sales as a proxy for investment opportunities, suggesting that our relationship variables are not proxies for investment opportunities. We also include the industry mean profits and mean standard deviation of profits as defined in Section III. These coefficients have the correct sign and are statistically significant, but they do not change our estimates of the coefficients on the relationship variables (see Table VIII, column 5).

As a further check, we include in the regression proxies for standard industry practice in regard to paying late. If most firms pay late, paying late must not be very costly. Therefore the fraction of firms in the two-digit SIC industry paying more than 50 percent late is an inverse measure of the penalty for paying late. The Late Payment Stretch in the two-digit SIC industry is a second measure of the net benefit of paying late. Though we lose a number of observations when we include these two variables, the relationship coefficients are not significantly altered by these additions. Two of the relationship coefficients are higher and one is lower.[27] Thus the regression is robust to proxies for the costs and benefits of paying late.

The extent to which a firm takes cash discounts for early payment is an (inverse) measure of credit availability and should be driven by the same factors that make a firm avoid penalties for late payments. Thus, the regression with "discounts taken" as the dependent variable should be viewed as a test of the robustness of our results. We expect the coefficients on the relationship variables to have the opposite sign in comparison to the previous regression. The results are reported in Table IX and confirm our earlier results. Stronger relationships are correlated with greater credit availability. The only additional point to note in these regressions is that in column 6 of Table IX, we include the implicit interest rate calculated from standard terms

[27] We lose observations because we only include firms in industries with at least 10 firms. This restriction ensures our estimates of medians are reasonable. Ideally, we should define the Late Payment Stretch as the difference in DPO between firms paying 100 percent of their credits late and those paying 0 percent late. We use the definition in Section IV.*A* so as to get sufficient observations to estimate medians precisely in each group.

Table IX

Credit Availability and the Role of Relationships
Dependent Variable: Offered Discounts Taken by the Firm (%)

The dependent variable is the percentage of early payment discounts that are taken by the firm. The coefficient estimates are from a tobit regression with two-sided censoring. The dependent variable is censored at 0.0 and 1.0. Asymptotic standard errors are in parenthesis. Except for column 6, each regression also includes seven industry dummies based on the one-digit SIC codes.

Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(book value of assets)	6.795*	6.554*	7.650*	6.747*	6.541*	8.549*	6.750*
	(1.679)	(1.691)	(1.670)	(1.778)	(1.678)	(2.718)	(1.751)
Profits/book assets	7.485*	7.657*	7.706*	6.493**	7.593*	14.022**	8.144*
	(2.399)	(2.419)	(2.428)	(2.624)	(2.373)	(5.601)	(2.600)
Debt from institutions/ book assets	−14.178**	−14.326**	−14.825**	−16.020**	−15.736**	−34.934**	−14.326**
	(7.161)	(7.168)	(7.217)	(7.437)	(7.184)	(13.736)	(7.353)
Firm is a corporation (= 1 if yes)	−9.913***	−9.844***	−9.811***	−12.751**	−8.831	−29.610*	−8.643
	(5.697)	(5.696)	(5.732)	(6.085)	(5.713)	(11.277)	(5.968)
Firm age (in years)[a]	0.900*	0.911*	6.612***	0.972*	0.899*	1.085*	1.003*
	(0.246)	(0.246)	(3.474)	(0.271)	(0.246)	(0.392)	(0.254)
Length of longest relationship (in years)[a]	0.904*	0.901*	17.154*	0.878*	0.884*	0.598***	0.841*
	(0.219)	(0.219)	(3.689)	(0.227)	(0.218)	(0.320)	(0.224)
Debt from financial service provider (%)	5.655		4.415	5.875	5.919	10.517	5.650
	(5.667)		(5.705)	(5.969)	(5.660)	(9.670)	(5.817)
Debt from financial service provider (only one service) (%)		10.159					
		(6.939)					
Debt from financial service provider (multiple services) (%)		1.539					
		(6.731)					
Number of institutions from which firm borrows	−8.889*	−8.790**	−9.754*	−9.486*	−8.789*	−11.192**	−7.505**
	(3.152)	(3.152)	(3.165)	(3.331)	(3.155)	(5.194)	(3.272)
Herfindahl index for financial institutions (1, 2, or 3)	14.996*	15.111*	14.608*	15.658*	14.170*	16.202*	14.449*
	(3.033)	(3.035)	(3.050)	(3.189)	(3.038)	(5.140)	(3.136)
Sales growth (1986–1987)				1.254			
				(3.744)			

Table IX—*Continued*

Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mean 1987 gross profits/assets ratio in two-digit SIC industry					78.292** (33.676)		
Mean σ(gross profits/assets) between 1983–87 in two-digit SIC industry					−69.090** (27.350)		
Interest rate implied by trade credit terms in two-digit SIC industry						−0.017 (0.063)	
Discount stretch (in days)[b]							−0.036 (0.190)
Fraction of firms taking $\geq$ 90% of early payments discounts							146.60* (29.188)
Number of observations	1500	1500	1500	1362	1500	545	1328
χ^2	194.1	195.3	185.7	183.1	202.8	76.7	216.5
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

[a] We replace length of relationship and firm age by the natural log of one plus the length of relationship and firm age in column 3. Thus the coefficient measures the change in the interest rate due to a one percent increase in the firm's age or the length of its longest relationship.

[b] For each two-digit SIC industry, the median DPO is obtained for firms taking more than 90 percent of discounts offered. This is subtracted from the DPO for firms taking less than 10 percent of discounts offered to obtain the discount stretch.

* Significant at the 1 percent level.

** Significant at the 5 percent level.

*** Significant at the 10 percent level.

of trade credit for the two-digit industry to which the firm belongs.[28] We lose two thirds of our observations, so these results must be interpreted with caution. We find that higher implicit rates have almost no effect on the percent of discounts taken. The coefficient is actually negative, but its magnitude is tiny. That the implicit interest rate has so little effect may imply that we are measuring it very inaccurately or that trade credit costs so much more than other sources that managers do not use it unless they have no other source of capital, an assumption implicit in our analysis.[29]

Clearly, our evidence that trade creditors lend when institutional lenders do not suggests that they have collateral, incentives related to the product they are selling, sources of leverage over the firm, or information that the institutions do not possess. Is it then possible that our relationship variables identify firms whose strong supplier relationship—and hence cheap trade credit—substitute for bank relationships and bank credit? For instance, suppliers may allow younger firms greater leeway in stretching out their trade credit repayments. If so, the negative correlation between age (or length of relationship) and the extent of late payments simply reflects the fact that the implicit cost of trade credit is lower for young firms. The data in Table VII, Panel B, however, do not support this explanation. The median stretch (as measured from the due date) for the youngest 10 percent of the firms is -5.86 days compared to a median stretch of -0.72 days for the oldest ten percent of the firms. Similarly, the median stretch for the smallest 10 percent of the firm is -17.91 days compared to a median stretch of 2.85 days for the largest 10 percent of the firms (see Table VII, Panel A). If, as suggested in Section IV.*B*, trade credit terms are uniform in an industry, it would imply that firms borrowing the most against trade credit are allowed considerably

[28] The terms were obtained from Dunn and Bradstreet's *Handbook of Credit and Collection* (1970). We obtained standard terms for 46 four-digit SIC industries which translated into 19 two-digit industries. We calculated the implicit interest rate assuming that the credit period began on the last day the discount could be used and continued till the day the payment was due (this assumption is consistent with our finding that the stretch in the retail and wholesale industry is somewhat smaller than the 20 days implied by the 2–10–30 rule). Whenever we had different terms for the same two-digit industry, we took a simple mean of the calculated implicit interest rates. The largest implicit interest rate (without considering those with cash terms where the due date and the discount date were the same) was 348 percent, and the lowest was 15 percent. The mean rate was 70 percent. The most common terms were 2–10–30, which were offered in 23 of the 46 four-digit SIC industries.

[29] Does the fact that firms borrow against trade credit even when the average implicit interest rate on the credit is 70 percent imply that the rate of return on the firm's marginal projects is higher than 70 percent? Clearly not. But as the following example shows, project indivisibility or nonconvexity is enough to rationalize the use of expensive trade credit. Consider a firm which has a $100,000 investment in equipment which will be liquidated at a fire sale price of $90,000 (see Shleifer and Vishny (1992)) if creditors get control rights over the firm. Further, assume a coupon payment of $5,000 is coming due. If the firm has no money to make this payment and no institution will lend more, it may borrow the $5,000 against trade credit to make the payment, in order to avoid the potential loss of $10,000 if creditors gain control. Even though the potential loss from project liquidation is only 10 percent of its value, the rate of return on the usage of trade credit is enormous. A similar point can be made for project initiations.

less stretch, and consequently pay considerably higher implicit interest rates on their trade credit borrowing. By contrast, interest rates on institutional loans are relatively less influenced by age and size (see Tables IV and X).

There is further evidence that trade credit is not meant to be a cheap substitute for medium-term financing. It is the practice in some industries for suppliers to finance buyers. The large volume of loans from nonfinancial firms in those industries is evidence of this. If supplier financing is explicitly intended to be medium term, we would not expect trade credit to be offered with discounts for early payment. This is indeed the case. Firms which have their largest source of loans from other nonfinancial firms were offered, on average, discounts with only 22.7 percent of their trade credit. By contrast, other firms are offered discounts on 32.9 percent of their trade credit. The difference in means is significant at the 5 percent level ($t = 2.4$). While trade credit may be the only source of finance when firms are young, the evidence that firms borrowing most on trade credit pay relatively the highest rates for it, and the evidence that suppliers who want to offer medium term credit offer explicit loans rather than trade credit, suggests that firms use trade credit out of necessity rather than choice.

A final possibility is that the relationship variables somehow proxy for firms in distress. If small, highly leveraged, and floundering firms are cut off by institutions (thus cutting short their relationship), and are forced to use trade credit, we would find a spurious negative correlation between the length of the relationship and the usage of trade credit. Are the firms using lots of trade credit necessarily distressed? The median firm repays 10 percent of its trade credits late. For firms paying more than the median late, the

Table X

Average Interest Rates on Most Recent Loan

The table contains the average interest rate on the firm's most recent loan categorized by the firm's book value of assets, the percent of trade credits paid late, and the percentage of early payments taken. The first row of the table contains the smallest firms, the firms that pay the largest percent of their trade credits late, and the firms that take advantage of the fewest early payment discounts.

	Percentiles Based on the Book Value of Assets	Percentiles Based on the Percent of Trade Credits Paid Late	Percentiles Based on the Percent of Early Payment Discounts Taken
0–10%	12.0	11.4	11.2
10–25%	11.5	11.9	11.1
25–50%	11.5	11.4	11.1
50–75%	10.8	11.2	10.8[a]
75–90%	10.5	10.7	10.8
90–100%	10.1	10.4	10.8

[a] Over 25 percent of the firms take all of the early payment discounts that are offered. Thus the groups 50–70 percent, 75–90 percent, and 90–100 percent are not distinct. Thus 10.8 percent is the average interest rate for firms taking more than the median percent of the early discounts which they are offered.

mean asset size is $1.17 million, the mean profitability (as a fraction of assets) is 0.35, the mean sales growth is 0.25, and the mean debt-to-assets ratio is 0.32. This compares to a mean asset size of $1.21 million, mean profitability of 0.44, sales growth of 0.19, and indebtedness of 0.26 for firms below the median. Only the debt levels are statistically different, though this may not reflect distress but simply that firms paying late are investing more (and hence are less profitable either because projects have not come on line or because they have greater depreciation tax shields) and using external financing. We also examine the difference between the age of the firm and the length of the longest relationship. If highly indebted firms are being cut off by their banks, leading to the spurious correlation suggested above, we should find the difference to be the highest in the case of the most indebted firms. Instead, we find that the longest relationship for highly indebted firms—firms with institutional debt above the median—is 1.3 years longer (relative to their age) than for firms with institutional debt below the median. Finally, Table X shows the average interest rate charged on the firm's most recent institutional loan. Firms using the most trade credit do not pay substantially more for their loans, suggesting indeed that we are measuring some form of credit rationing and not some spurious correlation arising from distress.

V. Discussion and Conclusion

We began our empirical investigation by noting that borrowing by small firms is highly concentrated. Moreover, small firms borrow a significant fraction of their debt from lenders who provide them informationally intensive financial services. Are there benefits to concentrating borrowing and building relationships with a few lenders or is such concentrated borrowing costly? Our analysis indicates the former.

We find a small effect of relationships on the price charged by lenders. The length of an institution's relationship with the firm seems to have little impact on the rate. Similarly, the rate charged is insignificantly lower when the lender provides the firm financial services. We find that firms that borrow from multiple banks are charged a significantly higher rate. There are a number of potential explanations of this effect, other than that multiple sourcing weakens relationships, but we do not find strong support for any of them.[30]

It does not appear that the lack of explanatory power occurs because our proxies for the strength of relationships are faulty. Using similar proxies, we

[30] Conversations with bankers provide some casual support for the "weakening of relationships" explanation. One banker said that he invariably tries to be the sole lender. If the firm asking for a loan has a prior relationship with another bank, he usually insists on "taking out" the prior bank with part of the new loan. Being the sole lender improves his ability to control the borrower's actions. Another banker echoes these feelings, adding that firms tend to change banks primarily when their existing bank reaches its legal lending limits. In such cases, a firm occasionally insists on maintaining token ties with its old bank. He also feels that some small business owners have "outsize egos," leading them to believe that their firms are big enough to warrant multiple banking relationships, even though it is a costly practice.

find stronger effects of relationships on the availability of financing. The empirical results suggest that the availability of finance from institutions increases as the firm spends more time in a relationship, as it increases ties to a lender by expanding the number of financial services it buys from it, and as it concentrates its borrowing with the lender.

The results from the previous section rule out the possibility that relationships have no value. They also indicate that our proxies are indeed capturing some aspects of relationships. There are at least two theoretical explanations as to why the burden of adjustment to strong relationships falls on the availability of credit more than it does on price. First, if Stiglitz-Weiss credit rationing is indeed taking place, the firm's marginal returns from investment may be much higher than the price of credit. Therefore, if offered a choice, firms would prefer more, rather than cheaper, credit. Unfortunately, peripheral evidence on this hypothesis is decidedly mixed. When the SBA Survey asked firms about the most important characteristic of financial institutions, "interest rates and prices offered" was the most frequent response (27.3 percent) while "a willingness to extend financing" was in second place (23.8 percent). However, when asked about the least important characteristic of financial institutions, "a willingness to extend financing" was the least common response (5.6 percent) while "interest rates and prices offered" came next (10.8 percent).

The other theoretical explanation is that while the relationship reduces the lender's expected cost, it also increases its informational monopoly, so that cost reductions are not passed on to the firm. We cannot distinguish between these two possibilities.

The different effects on price and quantity may also stem from the organizational structure of lending institutions. In order to maintain adequate checks and balances in their business, financial institutions have fairly specific guidelines for loan pricing. It would be difficult, and perhaps defeat their purpose, for the institution to set these guidelines such that the loan officer's "soft" information about the firm can be embedded in the price. Given this structure, it may be much easier for the loan officer to use her knowledge to influence the loan amount and whether the loan is made at all, rather than the price.

Our study also throws additional light on another important public policy issue. A bank may have economic value because it screens out poor credits. But once the public credit market knows which firms are good (by observing firms that have had a long relationship), there is no externality imposed on the firms if the bank fails or is forced to contract its lending. On the other hand, if a bank generates substantial durable and nontransferable private information during the course of a relationship, there may be significant externalities when it fails or reduces lending commitments, because others cannot easily step into the breach (see Bernanke (1983)). Slovin, Sushka, and Polonchek (1993) provide evidence that banks may, in fact, serve as repositories of private information. They find that the impending insolvency of Continental Illinois Bank had negative effects and the FDIC rescue had positive effects on client firm prices. Our study adds to theirs by detailing the mechanisms through which the bank may acquire information about the firm, and how it passes on the benefits of this more intense monitoring back to the firm.[31] The public policy implication is that regulators should factor in the informational capital that will be destroyed when deciding whether to save a bank from liquidation.

Perhaps the most interesting conclusion of our study is that the apparent concentration of borrowing and the purchasing of financial services does not seem to make small firms worse off. Small firms may voluntarily choose to concentrate their borrowing so as to improve the availability of financing. Furthermore, we find that firms in areas where there are few bank-like institutions are less likely to be rationed. This accords with the notion in Mayer (1988) and Rajan (1992) that increased competition in financial markets reduces the value of relationships because it prevents a financial institution from reaping the rewards of helping the firm at an early stage. The policy implication is that these firms may best be helped if lenders can make their claims to the firm's future profits explicit; for instance, regulations prohibiting banks from holding equity could be weakened so that banks have an explicit long-term interest in the firms to which they lend.

[31] On its own, our study cannot fully resolve whether the information generated in a relationship is private or public. It is possible that the length of the relationship is a significant determinant of the availability of credit, not because the creditor has accumulated private information about the firm, but because creditors attempt to keep the business of their best credits as long as possible. The length of the relationship may then be a publicly available proxy, similar to the age of the firm, of a firm's creditworthiness. It is, however, harder to explain why availability increases as creditors come closer—where "closeness" is measured by the number of nonfinancial services they offer the firm—unless we accept that some private information is generated via these services. None of these services are so specialized or sophisticated that only "high-quality" managers would think of using them. Only a few of these services (banker's acceptances and letters of credit) force the bank to take on credit risk, and these commitments are usually short term and well secured so that the credit risk is minimal. It is hard to think of how the provision of these services could be a public signal of quality. It is, however, possible that the provision of these services helps tie the firm to its creditor in the long run, making the creditor more willing to extend funds.

REFERENCES

Ausubel, Lawrence M., 1992, Rigidity and asymmetric adjustment of bank interest rates, Mimeo, Kellogg Graduate School, Northwestern University.

Berger, Allen, and Gregory Udell, 1992, Small firms, commercial lines of credit, and collateral, Mimeo, Board of Governors of the Federal Reserve System.

Bernanke, Ben, 1983, Non-monetary effects of the financial crisis in the propagation of the Great Depression, *American Economic Review* 73, 257–276.

Blackwell, N., and Anthony Santomero, 1982, Bank credit rationing and customer relation, *Journal of Monetary Economics* 9, 121–129.

Brown, Charles, James Hamilton, and James Medoff, 1990, *Employers Large and Small* (Harvard University Press, Cambridge, Mass.).

Campbell, Tim, and William Kracaw, 1980, Information production, market signalling, and the theory of intermediation, *Journal of Finance* 35, 863–882.

Dennis, William I., William C. Dunkelberg, and Jeffrey S. Van Hulle, 1988, *Small Business and Banks: The United States* (N.F.I.B. Foundation, Washington, D.C.).

Diamond, Doug, 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 51, 393–414.

———, 1991, Monitoring and reputation: the choice between bank loans and directly placed debt, *Journal of Political Economy* 99, 688–721.

Dun and Bradstreet, 1970, *Handbook of Credit Terms* (Dun and Bradstreet, New York).

Elliehausen, Gregory, and John Wolken, 1992, The use of trade credit by small businesses, Mimeo, Board of Governors of the Federal Reserve System.

Fama, Eugene, 1985, What's different about banks?, *Journal of Monetary Economics* 15, 29–36.

Fazzari, Steve, Glenn Hubbard, and Bruce Petersen, 1988, Investment and finance reconsidered, *Brookings Papers on Economic Activity* 141–195.

Greenbaum, Stuart, George Kanatas, and Itzhak Venezia, 1989, Equilibrium loan pricing under the bank client relationship, *Journal of Banking and Finance* 13, 221–235.

Haubrich, Joseph, 1989, Financial intermediation, delegated monitoring, and long-term relationships, *Journal of Banking and Finance* 13, 9–20.

Hodgman, Donald, 1963, *Commercial Bank Loan and Investment Policy* (Bureau of Business and Economic Research, University of Illinois, Chicago).

Hoshi, Takeo, Anil Kashyap, and David Scharfstein, 1990a, Bank monitoring and investment: Evidence from the changing structure of Japanese corporate banking relationships, in R. Glenn Hubbard, ed.: *Asymmetric Information, Corporate Finance and Investment* (University of Chicago Press, Chicago).

———, 1990b, The role of banks in reducing the costs of financial distress in Japan, *Journal of Financial Economics* 27, 67–88.

———, 1991, Corporate structure, liquidity and investment: evidence from Japanese industrial groups, *Quarterly Journal of Economics* 106, 33–60.

James, Christopher, 1987, Some evidence on the uniqueness of bank loans, *Journal of Financial Economics* 19, 217–235.

——— and Peggy Wier, 1990, Borrowing relationships, intermediation, and the cost of issuing public securities, *Journal of Financial Economics* 28, 149–171.

Leland, Hayne, and David Pyle, 1977, Information asymmetries, financial structure, and financial intermediaries, *Journal of Finance* 32, 371–387.

Lummer, Scott, and John McConnell, 1989, Further evidence on the bank lending process and the capital market response to bank loan agreements, *Journal of Financial Economics* 25, 99–122.

Mayer, Colin, 1988, New issues in corporate finance, *European Economic Review* 32, 1167–1189.

Mian, Shehzad, and Clifford Smith, 1992, Accounts receivable management policy: Theory and evidence, *Journal of Finance* 47, 169–200.

Petersen, Mitchell, and Raghuram Rajan, 1993, The effect of credit market concentration on lending relationships, Working paper, University of Chicago.

Rajan, Raghuram, 1992, Insiders and outsiders: the choice between informed and arm's length debt, *Journal of Finance* 47, 1367–1400.

Roosa, Robert, 1951, Interest rates and the central bank, in *Money Trade and Economic Growth: Essays in Honor of John Henry Williams* (Macmillan, New York).

Sharpe, Steven, 1990, Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships, *Journal of Finance* 45, 1069–1087.

Shleifer, Andrei, and Robert Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance* 47, 1343–1366.

Shockley, Richard, and Anjan Thakor, 1992, Information content of commitments to lend in the future: Theory and evidence on the gains from relationship banking, Discussion Paper No. 523, Indiana University.

Slovin, Myron B., Marie E. Sushka, and John A. Polonchek, 1993, The value of bank durability: Borrowers as bank stakeholders, *Journal of Finance* 48, 247–266.

Smith, Janet K., 1987, Trade credit and information asymmetries, *Journal of Finance* 42, 863–872.

Stiglitz, Joseph, and Andrew Weiss, 1981, Credit rationing in markets with imperfect information, *American Economic Review* 71, 393–410.

June, 1994

THE EFFECT OF CREDIT MARKET COMPETITION ON LENDING RELATIONSHIPS*

Mitchell A. Petersen and Raghuram G. Rajan

ABSTRACT

This paper provides a simple framework showing that the extent of competition in credit markets is important in determining the value of lending relationships. Creditors are more likely to finance credit constrained firms when credit markets are concentrated because it is easier for these creditors to internalize the benefits of assisting the firms. The paper offers evidence from small business data in support of this hypothesis.

* We thank Alan Berger, Judith Chevalier, Constance Dunham, Mark Flannery, Michael Gibson, Anne Grøn, Oliver Hart, Steven Kaplan, Bob McDonald, George Pennacchi, Canice Prendergast, Rafael Repullo, Ivo Welch, Lawrence White, and John Wolken for valuable comments on a previous draft. We are again grateful to John Wolken for making a special effort to provide us with some of the data. The editor, Andrei Shleifer, and two anonymous referees made suggestions that improved the paper significantly. We also thank workshop participants at Brigham Young University, the University of California at Los Angeles, Carnegie Mellon University, the Center for Economic Policy Research Conference on Financial Regulation at Toulouse, the Conference on Industrial Organization and Finance at San Sebastian, the University of Chicago, Columbia University, the Federal Reserve Board, the Federal Reserve Banks of Atlanta and Minneapolis, Harvard Business School, Indiana University, London School of Economics, Massachusetts Institute of Technology, Northwestern University, Southern Methodist University, Stanford University, Virginia Polytech, University of Wisconsin and the Utah Winter Finance Conference for their insights.

Is it possible for firms facing competitive credit markets to form strong ties to particular creditors? Does the benefit to a firm of forming such relationships diminish when credit markets get more competitive? There is a theoretical reason for believing that credit market competition may be inimical to the formation of mutually beneficial relationships between firms and specific creditors. When a firm is young or distressed, the potential for future cashflows from its projects may be high, while the actual cash it generates is low. When evaluating the credit-worthiness of the firm, a creditor should take into account the stream of future profits the firm may generate. When the credit market is competitive and creditors cannot hold equity claims, the lender cannot expect to share in the future surplus of the firm. She is constrained to break even on a period by period basis because she would drive away business if she charged a rate above the competitive one. Since uncertainty about a firm's prospects is high when the firm is young or distressed, creditors in a competitive market may be forced to charge a high interest rate until the uncertainty is resolved. This can be extremely distortionary to the firm's incentives and may, in fact, result in the firm not receiving credit at all. A monopolistic creditor, on the other hand, shares in the future surplus generated by the firm through the future rents she is able to extract. She can backload interest payments over time, subsidizing the firm when young or distressed and extracting rents later. Consequently, she may be more willing to offer credit than a similarly placed lender in a competitive market. In other words, credit market competition imposes constraints on the ability of the firm and creditor to intertemporally share surplus. This makes lending relationships less valuable to a firm because it cannot expect to get help when most in need.

The argument that relationships and competition are incompatible recurs in other sub-disciplines in economics.[1] For instance, labor economists claim that a firm is more reluctant to invest in training workers in a competitive labor market unless they post a bond, since workers can threaten to quit and demand a competitive salary once they are trained [Becker 1975]. While there is anecdotal evidence in the financial press suggesting this kind of phenomenon is a reality, there is little formal supporting evidence in the academic literature.[2] This is partly because periods of increased competition are, in general, also accompanied by other changes which may make relationships less valuable. It is thus difficult to disentangle the effects. For example, Hoshi, Kashyap and Scharfstein [1990] document that high quality Japanese firms moved away from their banks when the domestic bond markets were liberalized. This movement may reflect the adverse effects of competition on relationships. Alternatively this movement may simply reflect a reshuffling of borrowing by firms that previously did not have access to public markets. The ideal test would examine the effects of competition on relationships in a situation where firms are not simultaneously faced with a change in the sources of capital available to them.

In this paper, we focus our analysis on small businesses in the United States. Such firms concentrate their external borrowing from banks [see Petersen and Rajan 1994] and therefore the confounding effect of the portfolio choice problem discussed above is minor. In different regions of the country, the degree of competition between banks varies, partly because of restrictions on bank entry and branching, partly because there is only so much banking business a local market can support. Thus, we can isolate the effects of competition on relationships in this sample.

In the next section, we present a simple model which highlights the distinction between competitive and concentrated credit markets. In section II, we describe the data and the empirical tests. We find that significantly more young firms obtain external financing in concentrated markets than in competitive markets. Underlying differences in firm quality across markets do not appear to explain the differential access to capital. As the theory suggests, creditors smooth interest rates intertemporally in more concentrated markets, which may explain why they are able to provide more finance. We conclude in section III with policy implications and suggestions for future research.

I. The Model

A. Agents and Investment Opportunities

Assume a risk neutral world where there are two types of agents looking for finance; good entrepreneurs and bad entrepreneurs. At date 0, good entrepreneurs can choose either a safe project or a risky project. A safe project pays out S_1 at date 1 when amount I_0 is invested in it. Furthermore, when the project concludes the good entrepreneur will be able to invest I_{1S} in another safe project which returns S_2. Alternatively, if the good entrepreneur chooses to gamble at date 0, he can invest I_0 in the risky project. At date 1, the risky project may succeed and pay out R_1 with probability p, or it may fail and pay 0 with probability (1-p). If it succeeds, he can invest I_{1R} in a safe project which pays R_2 at date 2. In contrast to good entrepreneurs, the projects of bad entrepreneurs are doomed to fail and return nothing at date 1.[3]

We also make the following assumptions:

(A.1)–(A.4) [illegible]

A.1 is simply that the safe project has positive net present value (NPV) in this risk neutral world, while A.2 implies the risky project has negative NPV. A.3 implies that the future positive NPV project has the same expected returns and investment, regardless of whether the safe or risky project is chosen at date 0. Finally, A.4 implies that the revenue from the date-0 project is insufficient to finance the project at date 1.

B. Finance.

Financial institutions are the only source of external finance in this market. For simplicity, we call these institutions banks. Agents know whether they are good or bad entrepreneurs. At date 0, banks only know that a fraction θ of the agents that demand finance are good entrepreneurs. Thus θ is a measure of the ex-ante credit quality of the agents. A number of studies suggest that a lending bank learns a lot about the kind of entrepreneur it is dealing with over the course of its relationship with the entrepreneur [for example, see the evidence in Lummer and McConnell 1989, and Petersen and Rajan 1994]. Therefore, we assume that at date 1 the bank becomes fully informed about the kind of agent with whom it is dealing. We also assume that for regulatory reasons banks can only hold debt claims, that is, contracts which require a fixed repayment by the borrower.[4] Due to the difficulty of describing investments or the character of the entrepreneur to courts, contracts cannot be made contingent on the project taken or the type of agent.[5] Finally, the bank can charge a rate such that its expected return on loans is M. M is a measure of the market power the bank has. In this risk neutral world where the risk free rate is zero, we get perfect competition in the credit market when M equals 1, while a bank has market power when M is greater than 1. For simplicity, we assume[6]

(A.5) $$\frac{S_2}{I_{1S}} > M \geq 1 .$$

The borrowing process is as follows: the agent (the good or bad entrepreneur) goes to a bank and indicates how much he wants to borrow and for what maturity. The bank responds by quoting an interest rate which will give it an expected return less than or equal to M. If no interest rate gives the bank an expected return greater than or equal to its cost of funds (= 1), it turns down the loan. In what follows, we determine how the quality of firms getting financed, and the amount they pay for credit over time, varies with the bank's market power.

C. Solving the Model.

This problem is a straightforward application of Diamond [1989] and Stiglitz and Weiss [1981].[7] At date 0, good entrepreneurs cannot distinguish themselves from bad ones. Therefore, they can borrow only at a rate which compensates the bank for possible losses if the entrepreneur turns out to be bad. The higher interest rate can distort the entrepreneur's incentives and persuade him to choose the risky project. Thus adverse selection can cause moral hazard which in turn can lead to credit rationing.

To see this, first note that good entrepreneurs will try to reduce their own cost of borrowing by asking for terms that help expose the bad entrepreneurs. A bad entrepreneur will have no choice but to ask for the same terms at date 0. Since bad entrepreneurs have no new projects at date 1, the bank will not give them new money unless it has contracted to do so. Knowing this, good entrepreneurs will borrow as little as possible at date 0, so they can take advantage of the lower rate at date 1 when the bad entrepreneurs have been exposed. Therefore, a good entrepreneur will seek to borrow only I_0 at date 0. Without loss of generality, we assume that he proposes to repay amount D_1 at date 1, after which he will contract a new loan for any subsequent project.[8]

If the good entrepreneur chooses the safe project at date 0, at date 1 he must borrow

(1) $$I_{1S} - (S_1 - D_1).$$

He can now borrow at the riskfree full-information rate M.[9] If the good entrepreneur chooses the safe project at date 0, he expects

(2) $$\max\left[\, S_2 - M\left(I_{1S} - (S_1 - D_1)\right),\ 0 \,\right].$$

Similarly, his expected profit if he chooses the risky project is

(3) $$\max\{\, p\{R_2 - M[\, I_{1R} - (R_1 - D_1)\,]\},\ 0 \,\}.$$

Using assumptions A.3, A.4, and A.5, the good entrepreneur strictly prefers the safe project if

(4) $$\frac{S_1 - pR_1}{1-p} \geq D_1 .$$

For the bank to lend at date-0, two conditions must be satisfied. First the loan must be structured so that the good entrepreneur has the incentive to take the safe project. This is inequality (4). In addition, the bank must expect to recover its date-0 investment I_0. Taking into account the profit it can make by charging the maximum interest rate on loans made at date-1, this implies that the bank will lend only if

(5) $$D_1 \geq \frac{I_0}{2M} - \frac{M-1}{M}\left(I_{1S} - S_1\right).$$

Using (4) and (5), only entrepreneurs with credit quality greater than

(6) $$p^{C}(M) = \frac{I_0(1-p)}{M(S_1 - pR_1) - (M-1)(I_{1S} - S_1)(1-p)}$$

will get financed.

Result 1. As the market power M of the bank increases, firms with lower credit quality obtain finance.

The intuition is straightforward. As the market power of the bank increases, it can extract a larger share of the future surplus generated by the firm. This implicit equity stake in the firm enables it to set a lower interest rate for the initial project (compared to a more competitive situation). The surplus extracted in the future does not affect the firm's choice between projects (see equation (4)), but the lower initial rate give the firm an incentive to take the safe project. Consequently, firms of lower quality can be profitably financed.

The interest rate charged for funds lent at date 1 is M. But at date 0, the face value demanded, D_1 is bounded by:

(7) $$\left[\frac{I_0 M}{2}, \frac{S_1 - pR_1}{1-p}\right] \geq D_1 \geq \frac{I_0}{2M} - \frac{M-1}{M}(I_{1s} - S$$

The first argument of the minimum function is the limit set on the interest rate by the bank's market power, while the second is the limit set by moral hazard. The term after the second inequality is the bank's individual rationality condition. For the lowest quality firm financed by the bank with market power M, the binding upper limit on the interest rate is moral hazard. The lower binding constraint is the bank's individual rationality condition which depends on M. Therefore,

Result 2. The initial interest rate contracted by the lowest quality firm financed by a bank with market power M is lower than it would be if such a firm were to be financed by a bank with market power M′ where M > M′.

Finally, let t be the ratio of the face value charged per dollar lent at date 1 to the face value charged per dollar lent at date 0. This is the ratio of gross interest rates. Then for any initial credit quality θ and banks with market power M and M′ where M > M′,

(8) $$t(M) \geq t(M').$$

In a population of firms where θ has positive density everywhere on [0,1],

(9) $$\text{Average}(t(M)) > \text{Average}(t(M')).$$

Result 3. On average, the relative decline in demanded repayments as the firm gets older is lower when the bank has more market power.

D. Caveats: Contractual Remedies.

Our main point, similar to Townsend [1982] and Gertler [1992], is that multi-period state contingent contracts allow for more efficient contracting than single or multiperiod fixed payoff contracts. In the environment we analyze, backloaded state contingent interest payments are less distortionary than frontloaded fixed interest payments. A creditor with market power can convert the latter into the former, which is why efficiency increases with creditor market power. There are obvious caveats.

We have abstracted from a variety of contracts which could commit the firm to sharing surplus with the lender, even in a competitive environment. For example, we have assumed that regulators do not allow creditors to hold equity claims. However, equity contracts may not be feasible even if allowed, since dividend payments are voluntary. Diamond [1984] and Hart and Moore [1989] specify environments where firm's management will never voluntarily part with cash. One way to force management to pay out cash in these environments is through debt contracts. Another way that we emphasize above is for a supplier to have market power in an essential input (credit) which enables it to extract cash from the firm. Thus the implicit equity stake that market power confers may be feasible even when explicit equity stakes are not.

The reader may wonder why long term debt contracts between the firm and the bank in the competitive environment cannot replicate the same payments over time as those determined by market power. The reason, quite simply, is that the payments due on the long term debt do not -- and cannot -- vary with the project taken. The share of the surplus the bank gets by virtue of its market power does, however, vary. The firm has to pay more to the bank if the risky (but ex-post profitable) project succeeds. This reduces the owner's incentive to shift risk.[10] This also explains why contracts which require a high initial payment, and then commit the bank to relending at a lower-than-competitive rate will not work. These contracts have the effect of making the firm's payments conditional on the risky project succeeding lower and thus increase the firm's incentive to shift risk.

Within the class of debt contracts, however, complicated bonding contracts could commit the firm to sharing surplus. For example, the entrepreneur could commit to asking the bank for a loan at date 1 (giving it a 'first-right-of-refusal') at a pre-determined high rate.[11] There may be practical difficulties in enforcing such a contract, primarily because the successful firm in a competitive credit market can easily find ways of borrowing without violating the letter of the 'first-right-of-refusal' contract. For instance, the entrepreneur could take out a personal loan from another source, collateralized by the shares of the firm, or sell equity and use the proceeds to repay the high priced loan. A contract ruling out such contingencies would be costly to write and difficult to enforce. Another problem is that having bonded the entrepreneur, the bank has no incentive to offer good service. Finally, it is possible that a populist legal system may refuse to enforce contracts which appear (ex post) to be extortionary. Thus bilateral opportunism or political realities may make these contracts impractical.

We have also abstracted from the distortions to the firm's investment incentives created by the bank's market power. Even if these are large, results 1, 2, and 3 will continue to hold so long as the bank can contractually dispose of its ability to extract future rents, as for example, by signing a loan commitment contract. The contract would need a 'material adverse change' clause allowing the bank to deny credit to bad entrepreneurs. While such contracts are observed, opportunism may again make them impractical.

In summary, the theory is ambiguous on whether the problem we have described can be effectively contracted around, though casual empiricism (see note 2) suggests it cannot. The test of whether it is important, however, must be an empirical one.

II. Data and Empirical Investigation

A. Sample Description

The data in this study are obtained from the National Survey of Small Business Finances. The survey was conducted in 1988-89 under the guidance of the Board of Governors of the Federal Reserve System and the US Small Business Administration. It targeted non-financial, non-farm small businesses which were in operation as of December, 1987. Financial data was collected only for the last fiscal year. The firms in the survey are small -- fewer than 500 employees. The sample was stratified by census region (Northeast, North Central, South, and West), urban/rural location (whether the firm was located in a MSA), and by employment size (less than 50 employees, 50-100 employees, more than 100 employees). The stratification was done to ensure that large and rural firms were represented in the sample. The response rate was seventy to eighty percent, depending upon the section of the questionnaire considered.[12]

There are 3404 firms in the sample, of which 1875 are corporations (including S corporations) and 1529 are partnerships or sole proprietorships. In the overall sample, the mean firm size in terms of book assets is $1.05 million; the median size is $130,000. The firms have mean sales of $2.6 million and median sales of $300,000. These companies are also fairly young, having spent a median of 10 years under their current ownership. In comparison, firms in the largest decile of NYSE stocks have been listed for a median of at least 33 years. Nearly 90 percent of these firms are owner managed, 12 percent are owned by women and 7 percent by minorities. Over 19 percent of the sample consists of firms in the hotel and restaurant business. Another 13 percent are in building and construction, 7 percent manufacture intermediate products like chemicals, and 7 percent perform communication, electric, gas and sanitary services.

On average, firms in the sample have a debt to asset ratio of 0.33. Of the debt, approximately 80 percent is from institutions while the rest consists of loans from the owners (in the cases of partnerships or corporations) and the owner's family. Banks account for 80 percent of the institutional lending, and non- bank financial institutions about 15 percent.[13] On average, trade credit financing is 10 percent of total assets, or about a third of the size of debt.

B. Sources of Market Power

In order to proceed further, we must determine proxies for the bank's market power. Note that nothing in our analysis requires the market power to be present ex ante. One possibility then is that the bank obtains market power ex post from the private information it obtains about the firm during the course of the lending relationship. The problem with this kind of local monopoly is that it would dissipate as the firm gets older and better known by the credit market. A second problem is that firms who find it hard to get credit are likely to offer their banks a local monopoly, making the choice endogenous. Finally, it is empirically hard to quantify the extent of information asymmetries about the firm between the inside lender and the outside credit market.

Another source of market power is the spatial distribution of banks in the local market. Banks that are physically closer to the firm have lower costs of monitoring and transacting with the firm. These costs may be especially significant because the firms in our sample are small. If other banks are relatively far, close banks have considerable market power. The concentration of banks in the local geographical market is likely to be a proxy for how far competing lenders are likely to be from one other, and consequently how much market power any single lender has. Another way to think about this is that the search costs to a firm of finding a replacement lender who has the ability to deal with its specific needs are likely to be high when the local market has few lenders.

In what follows, we use the concentration of lenders in the local market as a measure of the lender's market power. Implicit in the analysis that follows is the following assumption: variations between local markets in the market power banks have because of their concentration in the local markets is much larger than variations between local markets in the amount of 'homemade' market power that firms voluntarily offer banks by giving them an information monopoly. If this were not true, our tests would be biased against finding any effect of spatial concentration on lending.

From the F.D.I.C. Summary of Deposit data, the SBA survey obtained the Herfindahl index of commercial bank deposit concentration for the county or Metropolitan Statistical Area where the firm is headquartered. The survey, however, reported only a broad categorization of competition in the local credit market. The survey reports whether the Herfindahl index is less than 0.1, between 0.1 and 0.18, or greater than 0.18. We refer to the first category as the most competitive credit market, and the last category as the most concentrated credit market.

In what follows, we take the concentration of the market for deposits to be a proxy for the concentration of the market for credit. This would be a good approximation if the firms in our sample largely borrow from local markets because of the prohibitive informational and transactional costs of going outside. While there is evidence in prior work [see Hannan 1991] of this, we have independent confirmation that credit markets for small firms are local.[14] The firms in our sample report the distance to their primary financial institution. Over half the firms are within 2 miles of their primary institution. Ninety percent of the firms are within 15 miles of their primary institution.

If the Herfindahl index for deposits is a good proxy for competition in the loan market, we would expect to find greater solicitation of new business by financial institutions in more competitive markets. For firms less than ten years old, 32 percent of the firms in the most competitive market have been approached by at least one financial institution seeking new business in the past year. Only 26 percent of the firms in the least competitive market were solicited. The difference in solicitation rates is marginally significant at the 13 percent level. The difference in solicitation rates is greater, both economically and statistically, for the older firms. For firms which are more than ten years old the solicitation rates are 31 percent in the most concentrated market and 46 percent in the most competitive market. This difference (statistically significant at the one percent level) is, therefore, greatest for older firms where the model argues competition is most destructive to relationships.

Solicitation of new financial business from the firms in this sample is once again a local phenomenon. Of the firms which were solicited for new business, over half report that the soliciting financial institution is within 3 miles of the firm. All this suggests that the Herfindahl index of deposits is a good measure of credit market competition. Before examining the data, it may be useful to restate our empirical predictions in terms of this proxy.

C. Empirical Implications

The empirical predictions of the model are the following:

a. Provided the distribution of firm qualities is similar in all markets and provided that in at least one market some firms do not obtain finance, relatively more firms should be able to obtain credit from financial institutions in areas where credit markets are more concentrated. Furthermore, the average quality of firms obtaining finance should be lower as the credit market becomes more concentrated.

b. Credit should be cheaper for the lowest quality firms in a concentrated credit market than if similar firms were to obtain credit in a more competitive market.

c. The cost of credit should fall faster as a firm ages in a competitive credit market than in a concentrated market.

D. Credit Market Competition and Firm Quality

The theory makes strong predictions about how borrowing will vary with firm age and market concentration. The sample we have is a cross section. However, by examining firms of different ages, we can make inferences about how borrowing changes over a firm's life.[15] The first prediction is that if the quality distribution of firms applying for credit is the same in the different markets, firms of lower average quality (and, consequently, relatively more firms) will be financed in more concentrated markets. We first describe the differences in quality of all the firms in the different markets in Table 1. Thus this table is based on both those firms which borrow from institutional sources and those which do not.

We divide the data into two subsamples -- firms younger than the median age of ten years and firms older than the median age. There are 296 firms in the most competitive credit markets and 2037 in the most concentrated markets. Firms in the two markets are approximately the same size as measured by the book value of total assets, with the median size of young firms being $102,000 and $103,000. There is greater disparity between the mean size of young firms ($863,000 in competitive markets and $569,000 in concentrated ones) but the difference is not statistically significant. Firms in the middle category (Herfindahl index between 0.10 and 0.18) are larger but not statistically so.

The young firms in the more competitive credit markets are somewhat more profitable. The ratio of gross profits (revenues less cost of goods sold) to assets is 2.78 in the competitive market and 1.88 in the most concentrated markets. The medians differ in similar ways. Moving on to operating profit ratios, again firms in the most competitive market are more profitable (a mean ratio of 1.14 versus 0.63), although the difference is no longer statistically different. The median ratios are even closer (0.17 versus 0.15). The rates of sales growth do not differ much across the markets (see Table I).

A firm's access to capital may depend upon the tangibility and/or the liquidity of its assets. The industry the firm is in should be a good proxy for this. As is typical for small firms, a significant fraction of our sample are in the retail trade and services industry. However, over ten percent of the firms are in construction and another ten percent of the firms are in manufacturing. The distribution of firms across industries is very similar in the most concentrated and the most competitive markets (see Figure I). The most competitive markets have relatively fewer firms in services and somewhat more in wholesale trade but the fractions of firms in the other industries are roughly equal across markets. In sum, the quality of firms in the most competitive market appears as good if not better than the quality of firms in the most concentrated market. Our results should be biased -- if at all -- towards finding more institutional financing in competitive areas.[16]

Finally, it is useful to ask whether the concentration of credit markets is correlated with the region of the country. This is important to ensure that we do not pick up the effects of a regional shock in what follows. A cross tabulation of the region in which the firm is located (Northeast, North Central, South, and West) and the level of competition in the local credit market (we do not report this in a table) indicates that firms in the southern and western regions are under-represented in the most competitive market and over-represented in the most concentrated markets. However, the magnitude of these differences is small. The correlations between the Herfindahl index and each of the four region dummies ranges from -0.12 to 0.15.

E. Credit Market Competition and the Availability of Finance

The theory in the previous section suggests that young firms are more likely to receive institutional finance in a concentrated credit market than in competitive one. Among the youngest half of our sample (firms which are less than ten years old), the firms in the most concentrated market are more likely to obtain capital from institutional sources (see Table II). Sixty five percent of the firms in the most concentrated market have institutional debt compared to only 55 percent of firms in the most competitive market, with the difference being significant at the 1 percent level. Ten year old firms may not be considered young. As Table II demonstrates, the difference in institutional financing is most extreme among firms which are four years old or less. Forty eight percent of such firms in the most competitive market have institutional debt compared to 65 percent of the firms in the most concentrated market. The difference is again statistically significant at the one percent level. The potential effects of a survival bias (see section II H) on our results will be smallest for the youngest firms in our sample. Yet in this part of our sample, the differential access to capital is the largest.

As firms grow older, the difference in the fraction of firms being financed in the two markets vanishes. For firms that are older than 10 years, approximately 61 percent use institutional finance, irrespective of the state of competition in credit markets (see Table II).

The differences in the fraction of young firms with institutional debt does not arise because firms in a concentrated market borrow less. The institutional debt to assets ratio for young firms which have institutional debt is slightly higher in concentrated markets -- 0.42 compared to 0.35. For firms older than ten years, the numbers are reversed. The institutional debt to asset ratio is 0.43 in the most competitive market and only 0.35 in the most concentrated market. As firms get older in the concentrated market, they borrow smaller amounts from these outside sources and rely more on equity and internal funds. This is consistent with our model. If borrowers in such markets are offered lower-than-competitive rates when young, this subsidy must be made up by charging them a higher-than-competitive rate when old. This will encourage them, at the margin, to substitute away from external borrowing when old.

In our discussion so far, we have implicitly assumed that the amount of institutional debt used is the amount of debt available to the firm. Such an assumption is, perhaps, defensible for the youngest firms which may have few internal sources of funds and many lucrative investment opportunities. For older firms it is less clear whether the amount of institutional debt is supply constrained (creditors do not want to lend more) or demand constrained (borrowers prefer internal sources). That older firms borrow less in concentrated markets does not necessarily indicate that these firms are more supply constrained. In order to explore this issue further, we turn to a better measure of credit availability.

F. An Alternative Measure of Credit Availability

If financial institutions limit the credit extended to a firm, the firm will borrow from more expensive non-institutional sources as long as the returns from its investments exceed the cost of funds from those sources. Correcting for investment opportunities, the amount of expensive borrowing should be a measure of how much the firm is rationed by the (cheaper) institutional sources -- provided the following conditions hold. First, the marginal cost of borrowing from the non-institutional source must exceed the marginal cost of available institutional credit. If it did not, the firm would turn to non-institutional sources of credit first. Second, the cost of borrowing from the non-institutional source should be relatively similar for firms within an identifiable class.

Most of the firms in our sample are offered trade credit -- short term financing which some suppliers provide with their goods and services. As reported in Table III, the median firm in all three kinds of markets obtains 90 percent of its purchases on credit. Another measure of the firm's use of trade credit is its Days Payable Outstanding (DPO) -- which is defined as 365 times the firm's accounts payables divided by its cost of goods sold. The median firm's DPO is about 15 days across the markets. We have data on the percentage of discounts for early payment which are taken by each firm. In general, discounts for early payment are substantial and are meant to encourage the firm to pay on time. For example, firms in the retail business refer to terms as the 10-2-30 rule. This is a discount of 2 percent if the bill is paid within 10 days and the full amount if paid in 30 days. A firm that passes up the discount and the bill when due pays an annualized rate of 44.6 percent, far higher than the highest interest rate charged to firms in our sample.[17]

Previous work indicates that discount terms are not specific to a firm, but are common practice throughout the industry [Elliehausen and Wolken, 1992, Petersen and Rajan, 1994]. Furthermore, although the decision to offer trade credit depends upon the firm's quality, the decision by trade creditors to offer early payment discounts does not. Approximately thirty-three percent of trade credits are accompanied by early payment discounts. This number does not vary with firm size or firm age. As these discounts and penalties are substantial and are industry- not firm-specific, the fraction of trade discounts not taken is a good proxy for the costly non-institutional credit source.[18]

By this measure, fewer firms appear to be credit constrained in concentrated markets than in competitive markets (see Table III). Only 19 percent of firms in concentrated markets take fewer than 10 percent of offered discounts, compared to 29 percent in the middle market and 33 percent in the most competitive market. Conversely, over 59 percent of firms in the most concentrated market take more than 90 percent of offered discounts, compared to 52 percent in the middle market and 50 percent in the most competitive market.

Clearly, these monotonic relationships are only suggestive since they do not control for firm quality or other factors which may be correlated with credit market concentration. To test the robustness of these results, we regress the percentage of trade discounts taken against measures of the firm's investment opportunities, its cashflow, measures of the strength of its lending relationships, and the state of competition in credit markets. Firms which take a larger percent of their early payment discounts should be less credit constrained. We include three measures of the firm's investment opportunities. Investment opportunities may depend on the firm's size -- the book value of its assets -- and the log of the firm's age (younger firms may have better opportunities). Since investment opportunities depend on the industry the firm is in, seven industry dummies are included as explanatory variables. Clearly, all these variables may also proxy for the credit quality of the firm.

The firm's internal cash flow is accounted for by including income after interest normalized by the firm's book value of assets. We include a dummy for whether the firm is a corporation or not, because credit rationing should be greater for firms with limited liability. An owner managed firm has a greater incentive to take on risky projects if it has limited liability. Petersen and Rajan [1994] find that the strength of relationships between firms and financial institutions are an important determinant of whether firms rely on trade credit financing. This is why we include the measures of relationships: the log of the length of the longest relationship the firm has had with a financial institution, the fraction of borrowing that comes from institutions that provide at least one significant financial service to the firm, and the number of institutions that account for more than 10 percent of the firm's borrowing.

Finally, we include an indicator if the firm is in the most concentrated credit market. If availability does not depend on the concentration of the credit market, the coefficient for this term should be zero.

The dependent variable in the regression should be the desired fraction of early payment discounts which the firm would like to take. Firms that are rationed by financial institutions will choose to borrow from trade creditors at the rates implicit in foregoing the early payment discount. In fact, they may wish to borrow more than is offered through their trade credits. Thus the desired percentage may be less than zero. Since firms cannot take less than 0 or more than 100 percent of the early payment discounts, the observed dependent variable is censored at 0 and 100 percent. In our sample, 60 percent of the firms are censored at 0 or 100. Estimating the model with ordinary least squares ignores this censoring, and consequently, estimates will be biased toward zero. We, therefore, estimate a tobit regression with two sided limits.

In Table IV we examine the determinants of the percent of offered discounts taken by the firm. The estimates in column I indicate that the investment and cashflow variables have the predicted sign. Older and larger firms have fewer investment opportunities and so can take more trade discounts by paying on time. If age and size proxy for credit quality, this result suggests that higher quality firms are less likely to be credit rationed. Profitable firms have more internal cash so they are less likely to use trade credit as a means of long term borrowing. Again, this may be a proxy for quality. Finally, because of their limited liability, corporations are more likely to take risky projects, which explains why they are more likely to be rationed.

The relationship variables also have the predicted signs. A long relationship with a financial institution increases the percentage of discounts taken, even holding the age of the firm constant. Borrowing a greater fraction from lenders who provide the firm services has a similar effect, although this effect is not statistically significant. Borrowing from multiple institutions makes relationships more diffused and increases the degree to which the firm is credit constrained.

The degree of concentration in financial markets enters in an economically and statistically significant way. A firm in the most concentrated market takes 17 percentage points more trade credit discounts than do firms in most competitive credit markets (t=2.9). This coefficient is forty percent of standard deviation of the percentage of discounts taken. In other words, after controlling for observable measures of credit demand and credit worthiness, we find firms in the most concentrated credit market are the least credit rationed.

An alternative explanation for our results is that credit constraints are less binding in small towns. Information about small businesses and their managers may be more available, or the pressure to repay debts may be greater, in rural areas. This should make credit rationing less severe. The correlation between our measure of credit market concentration and whether the firm is in a Metropolitan Statistical Area (MSA) is 0.44. The estimated coefficient for the indicator variable for MSA in Table IV column I is less than half that for the concentration indicator (β=-7.4) and statistically small (t=1.4). However, when we drop the indicator for SMSA from the estimation, the coefficient on the concentration indicator increases from 16.5 to 19.7. Thus the rural/urban divide has some effect in the predicted direction, but this does not account for the influence of concentration.

Another possibility is that the Herfindahl index is correlated with whether the firm is located in a state with a unit banking law. These states may have been especially affected by regional shocks -- for instance, to the oil or natural resources industries -- which in turn may have affected the entire regional economy.[19] Thus firms in competitive markets may be credit constrained not because the markets are competitive, but because they are in states affected by adverse regional shocks. The empirical facts are not, however, consistent with this alternative hypothesis. First, if the firms in competitive markets have been affected by regional shocks, then their profitability and sales should be affected by the shock. As is apparent in Table I, the firms in competitive markets have profit to assets ratios which are at least as large if not larger than the firms in the most concentrated markets. Sales growth for firms in the most concentrated market is the same as sales growth for firms in the most competitive market. Thus sales and profits of firms in these markets do not reflect adverse shocks.

A more direct test of this hypothesis is to control for the banking laws of the state in which each firm is located. This is done in the second column of Table IV. Firms in unit banking states are less credit constrained (β=9.3), although this coefficient is not statistically different from zero (t=0.9). More importantly, including a control for firms in unit banking states has only a marginal effect on the coefficient of the indicator variable for concentration. This coefficient actually rises from 16.5 to 17.1. Given the low correlation between the firm being located in a unit banking states and it being located in the most concentrated market (ρ=-0.07), this finding is not surprising.[20]

The evidence presented so far suggests that controlling for the observable measures of quality, firms in more concentrated credit markets are less credit constrained. The remaining portion of this section adds additional controls to the model to test the robustness of this finding and the accuracy of our assumptions. First, we have assumed that the firm's investment opportunities do not differ systematically with the concentration of the credit markets. The potential problem is that firms in more competitive credit markets may have greater investment opportunities and thus take fewer early payment discounts. As an additional control for investment opportunities we include the firm's sales growth, since intuitively firms with higher sales growth should also have more investment opportunities. The data, however, does not indicate that concentration proxies for differences in investment opportunities. The coefficient on the firm's sales growth is small and has the wrong sign. Firms with higher sales growth are slightly less credit constrained. In addition, the effect on the concentration coefficient is small. The coefficient on the Herfindahl index rises from 16.5 to 16.8 (see Table IV -- column III).

An alternative way of correcting for differences in industry profits, investment opportunities, and terms of trade credit in the industry is to include more detailed industry dummies. Instead of the seven industry dummies -- representing one-digit S.I.C. codes -- we include indicators for all two digit S.I.C. industries that account for more than 1 percent of the sample. We lose a number of degrees of freedom, and the concentration coefficient falls slightly to 14.5 (estimates not reported in table) but is still statistically significant (t=2.5).

The model in the previous section argues that the greater the lender's market power, the less credit constrained the borrower will be. So far, we have focussed on the differences between the most concentrated market and the other markets. But as Table II indicates, there seems to be a monotonic increase in the fraction of firms getting credit as we move across the three broad ranges for the Herfindahl. To test the monotonicity of this relationships in the trade credit regression, we add an additional dummy variable for the firms in the most competitive market. These estimates are reported in column IV of Table IV.

According to the model, the coefficient on the most competitive market dummy should be negative since these firms should be more credit constrained than the intermediate market which is the base. Instead the coefficient is positive. However, the coefficient is not statistically different from zero (t=1.1). Since our controls for relationships, such as the relationship length, depend upon the concentration of the credit market, we may be double counting. Interestingly, when we drop the relationship variables, the coefficient on the most competitive market drops to 3.7 (t=0.4). Thus we cannot distinguish the most competitive market from the middle market, although we can distinguish the most concentrated market from the other two.

Our inability to distinguish the estimates may be because the difference in credit availability between markets diminishes as firms become older and more established. Unfortunately the data are not discriminating enough for us to detect this phenomenon in the trade credit regressions. When we include both dummy variable for different levels of credit market competition and interactions with the firm's age (estimates not reported), the standard errors grow dramatically. However, we report means and medians of the percent of early payment discounts taken by firms in different markets in Table V. As expected from the model, we find that for young firms (age less than the median), the median discounts taken is related to the concentration of the market. The difference in medians is 10 percent higher in the middle market than in the most competitive market, and 30 percent higher in the most concentrated market than in the middle market. For the old firms, these differences vanish. The means follow a similar pattern, and unlike the medians, the differences in means are statistically significant.[21]

To summarize, we find that the amount firms borrow from financial institutions tends to increase with the concentration of the credit markets. When we examine a measure of availability of institutional finance rather than actual usage, we find that young firms appear less constrained in more concentrated markets. Regression estimates suggest that firms in the most concentrated market appear to be significantly less credit constrained than firms in the other two markets. The evidence is largely consistent with the theory proposed in section I. However to establish that the greater availability is because lenders in concentrated credit markets intertemporally smoothing interest rates, we now turn to analyze interest rates.

G. Cost of Capital Differences in Competitive and Concentrated Markets

Our model suggests that in markets where lenders have market power, they should charge a lower-than-competitive interest rate early in a firm's life, when problems of moral hazard and adverse selection are large. Later, they will compensate by demanding an interest rate above the competitive rate. For this reason, we expect the interest rate to fall more rapidly in competitive markets.

To test these predictions, we examine a subset of 1277 firms for which we have the interest rate charged on the firm's most recent loan. This is described in Table VI. The average interest rate is calculated for both the young (age less than 10 years) and old firms in the most concentrated and most competitive markets. Young firms pay higher average rates than old firms, but the pace at which the rate declines is larger in the most competitive market. In the most concentrated market the young firms pay 34 basis points more than old firms (t=2.17); however in the most competitive market this difference is 86 basis points. Notice that the interest rate starts higher in the competitive market and ends lower. This result is inconsistent with the argument that market concentration is a proxy for town size and the associated access to information. If this were true, rates should be uniformly lower in the concentrated market, not just for the youngest firms.

These results are only suggestive. We have made no adjustment for observable firm quality. According to the model in the previous section, lenders will lend money, on average, to lower quality firms in the more concentrated markets. Thus the difference in rates above between young firms in competitive and concentrated markets may understate the difference we should expect to find once we control for observable measures of quality.

To control for other factors which affect the interest rate a firm pays and to test the robustness of our findings, we estimate the loan rate as a function of the log of the firm's age and controls for observable firm and loan characteristics. In the regression results below we use the prime rate to control for changes in the underlying cost of capital. The prime rate includes both the risk free rate as well as a default premium for the bank's best customers. Since these small businesses are not the bank's best customers, they pay an additional default premium. We control for aggregate variations in this premium by including the difference between yield on corporate bonds rated BAA and the yield on ten year government bonds. We also include a term premium, defined as the yield on a government bond of the same maturity as the loan minus the treasury bill yield, to account for interest rate differences across different loan maturities. For floating rate loans this variable is set to zero.

To control for variation in the loan rate due to the characteristics of the loan we include dummies for whether it is a floating rate loan, for the kind of collateral offered, and for the type of lender making the loan. To control for variation in the loan rate due to the characteristics of the firm we include the firm's size (book value of assets), dummies for the firm's industry, and whether the firm is incorporated. We also include the relationship variables used by Petersen and Rajan [1994].

Since we expect firm age to have a different effect in competitive and concentrated markets, we estimate different age intercepts and slopes for each level of market concentration. The intercept measures the difference between a loan in the most concentrated and the most competitive market, when the firm is new. As Table VII shows, firms in the most concentrated market start out at an interest rate which is 129 basis points lower than in the most competitive market. Interest rates drop as the firm grows older, consistent with Diamond's [1989] notion that survival is a signal about the true quality of the borrower. The rate at which the loan rates drop, however, differs significantly across markets. The coefficient on age is over four times larger in the most competitive markets (Herfindahl index<0.10) than in the most concentrated markets (Herfindahl index>0.18). As a firm ages from new to the median age of ten years, its interest rate drops by 167 basis points in a competitive market. The interest rate drops only 36 basis points for firms in the most concentrated markets (see Figure II). In addition to being large in magnitudes, the difference in the slopes are also statistically significant (t=2.0). The fact that the slope of interest rates with age is flatter in concentrated markets is a direct implication of the our model.[22]

While it is interesting that interest rates start out lower for firms in the most concentrated area, it is not necessarily a prediction of our theory. Our theory only suggests that new firms of the lowest quality should obtain a lower loan rate in a concentrated market, it has nothing to say about the rate charged to the average new firm. The average loan rate in the concentrated market will be affected by both a higher markup charged to higher quality new firms (M>1) - which will tend to raise the average interest rate - and lower than competitive rates charged to low quality young firms - which will reduce it. The average effect depends on the distribution of firms. This may explain why the intercept is not monotonic, being even lower for firms in markets of intermediate concentration than in the most concentrated area (-1.70 and -1.29). But the slope coefficient in the intermediate market is not statistically or economically different from that for the most concentrated area (-0.135 vs. -0.154).

As discussed in section II.F., market concentration is correlated with the urban/rural location of the firm. As further evidence that concentration is not a proxy for the location of the firm, we re-estimate our rate regressions allowing the slopes to differ for firms in urban and rural locations rather than for firms in concentrated and competitive markets. Based on the results in the second column of Table VII, the rural/urban difference does not drive the results on concentration. The interest rate charged does fall with age, however, the rate of decline is very similar in rural and in urban markets. For firms in an MSA the coefficient on age is -0.240; for rural firms the coefficient is -0.154. The difference is not statistically significant (t=0.6).

Instead of market concentration being a proxy for the location of the firm (urban versus rural), the location of the firm is a proxy for market concentration.

Finally, we test for the effect of state banking laws on the rate. The inclusion of an indicator for unit banking states has little effect on the coefficient on age or the intercept in any of the markets (see Table VII column III). Our results appear to be robust; controlling for other firm and loan factors, the loan rates in the concentrated market start lower but do not fall as much with firm age as loan rates in the competitive market (see Figure II).

H. Survivorship Biases in Different Credit Markets

Interest rates slope downwards with firm age because good entrepreneurs survive while bad entrepreneurs go under. It is, however, important for us to ask if differences in the survival rate of bad entrepreneurs across markets could be responsible for the differences in estimated slopes across markets.

We assume in the model that bad types are all weeded out after one period. This ensures that the quality of surviving firms at date-1 is the same in all markets, even though lower quality firms are started in the concentrated market. An alternative possibility is that the greater availability of finance for low quality firms could make the attrition rate lower for bad firms in concentrated markets. Yet another possibility is that competition in the product markets may be fiercer in areas with competitive loan markets. Under the assumption that age proxies for unobservable elements of firm quality, we would find demanded interest rates fall less with firm age in concentrated markets.

Clearly, unobservable quality should be correlated with observable proxies for quality. Table I provides little evidence that surviving old firms in concentrated credit markets are of lower quality. The mean and median firm size and growth rates for old firms (firms greater than 10 years) is about the same in both competitive and concentrated markets though median profits -- but not mean profits -- are larger in the competitive markets. It is possible that most low quality firms have been weeded in all markets by the time firms are 10 years old. If differences in attrition rates are to explain the relative slopes of the interest rate with age, much of the difference in attrition rates should occur when firms are very young (because we have empirically documented a much steeper decline in interest rates with firm age for very young firms in competitive areas). We subdivide young firms into the very young (age less than 5 years) and the not so young (age greater than 5 but less than 10 years). If differential attrition takes place when firms are very young, the not so young firms should look much healthier in the competitive market. The evidence (not reported in table) does not support this. The not so young firms in the middle market dominate or tie with similar firms in competitive markets on all four dimensions of quality (median size, gross profits, net profits, and sales growth). The not so young firms in the concentrated market are the same size as those in competitive markets, have higher net profits but lower gross profits and sales growth. None of these differences are significant at standard levels.[23] Interestingly, size, which is the variable least likely to be mismeasured, is approximately the same at $ 121,000 across the markets. While we hesitate to dismiss differential survivorship as an explanation for the different slopes -- note, however, that it would not explain the differences in intercepts or the differences in availability -- we do not find evidence supporting it.

III. Discussion

Young firms in concentrated markets receive more institutional finance than do similar firms in competitive markets. As firms get old, the difference in the relative firms borrowing from institutions disappears. Young firms who get institutional loans are more indebted in concentrated markets than in competitive markets but this pattern reverses for older firms. Creditors seem to smooth interest rates over the life cycle of the firm in a concentrated market, charging a lower-than-competitive one when the firm is young and a higher-than-competitive rate when the firm is old. We now discuss possible interpretations of this result.

The results do not arise because firms in concentrated credit markets are very different from firms in competitive credit markets, since they persist even after we correct for the obvious firm and industry characteristics (see Tables IV and VII). In fact the sample of all firms in the different markets do not look very different (Table I). One possible explanation of our results is that banks in concentrated markets misprice their loans. But our results require them to be inconsistent in their mispricing, being excessively eager to lend to young, little known firms and excessively cautious in lending to older, established firms. While we cannot rule out this possibility, it is hard to think of plausible assumptions that would lead to such behavior. It is also possible that concentration is a proxy for other factors. It may be easier for small firms in rural markets to get credit, simply because lenders have better information about borrowers in a small community where news, and gossip, travel fast. The problem with this explanation is simply that ease of information acquisition cannot be the whole story, else firms in concentrated markets should enjoy uniformly lower rates independent of age. In addition, the results are not altered when we control directly for whether the firm is located in a rural area.

We think the most plausible explanation is the theoretical one provided in section I. Creditors in concentrated markets have an assurance of obtaining future surplus from the firm and consequently accept lower returns up front. This enables many more firms to be financed before the nominal interest rate rises to the point where the possibility of moral hazard forces the lender to cut off credit. Of course, banks in concentrated markets lend with the expectation that they will recover the initial subsidy via higher interest rates in the future. This explains why the surviving, older firms in concentrated markets pay higher nominal rates than surviving older firms in competitive markets. Finally, the natural response of older firms who have few investment opportunities and are faced with higher interest rates is to use less external finance. This would explain why older firms in concentrated markets rely on internal opposed to institutional finance.

There is a legitimate concern whether our findings are applicable to larger enterprises. The economic forces we have identified, however, are not special to small enterprises. While the main source of competition for a bank in the sample we examine is other banks, competition to fulfill the financing requirements of larger firms comes from arm's-length sources like the commercial paper, equity, and bond markets. The relative cost of borrowing from banks rather than markets is likely to be high for large, reputable firms [see Diamond, 1991, Rajan, 1992]. It is even less likely that banks will be able to retain the business of successful firms when competition is from arm's-length markets than when competition is from other banks. Therefore, a natural consequence of the opening up of arm's-length credit markets is that lending relationships deteriorate, young firms are more liquidity constrained, and banks are less willing to help their clients in financial distress.[24]

It is, however, easy to draw the wrong welfare implication from these findings. The absence of credit-market competition is only one way firms and creditors can share future surplus, and it is not costless. The higher interest rate paid by successful survivor firms in concentrated markets can be distortionary and force them to reduce their reliance on external funds. This may affect their investment and reduce their rate of growth. There are contractual ways of ensuring bilateral commitment between lender and borrower which may be less costly. For example, in addition to lending, the creditor could take an equity stake in the young or distressed firm.[25] While an equity claim would give the lender a share in the future surplus generated by the firm, explicit contracts (as opposed to implicit contracts enforced by the environment) have the added advantage of being voluntary so that firms whose incentives would be excessively distorted by sharing future surplus would not enter into them.

Of course, it may not always be possible to promote relationships via contracts. Equity contracts may not commit entrepreneurs to a credible payout policy. Creditors may not want to hold equity because it makes them 'soft' when dealing with errant borrowers [see Dewatripont and Tirole 1992]. Owners may be reluctant to give the bank equity because they may fear too much bank interference. A bank may be less inclined to monitor the firm on behalf of, and act in the interest of, other fixed claim holders (like employees and other creditors) if it also holds equity claims. A political concern is that banks may effectively bypass anti-trust laws if they build substantial equity holdings in similar companies. Clearly, it is an empirical matter whether these costs outweigh the benefits from stronger relationships, and our study has nothing to say about this.

It is interesting to speculate on the policy implications if the only way to promote relationships is by restricting credit-market competition. In the early stage of a country's economic growth when there are comparatively few established firms, the availability of finance is most important. It may be least inefficient to restrict inter-bank competition in order to achieve this.[26] But as many firms become established, the distortionary price for credit becomes a more important source of inefficiency. At this point, de-regulation and introducing more inter-bank competition, as well as opening up arm's-length markets may make the most economic sense.[27] Clearly, the period before an anticipated liberalization will be troubled as relationships breakdown before the benefits of competition kick in. But this may be unavoidable for an economy that wants to optimize the trade-off between the availability and the price of credit.

The general point the paper makes is that competition and long-term relationships are not necessarily compatible, whether the markets being analyzed are labor or capital markets. Reformers in relationship-based economies are not always right (as suggested by Mayer [1988]) if they think that adding a dose of competition to their systems will necessarily make it better off. Conversely, it may be equally wrong to expect firm-creditor or firm-worker relationships to be as strong or valuable in the competitive U.S. markets as in the cartelized markets elsewhere.

Northwestern University

University of Chicago

References

Becker, Gary S. *Human Capital* (Chicago, IL: University of Chicago Press, 1975).

Berger, Allen, and Gregory Udell, "Small Firms, Commercial Lines of Credit, and Collateral," Board of Governors of the Federal Reserve System, 1992.

Dewatripont, Mathias, and Jean Tirole, "A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence," C.E.P.R. Working Paper, 1992.

Diamond, Douglas, "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, LI (1984), 393-414.

________, "Reputation Acquisition in Debt Markets," *Journal of Political Economy*, XCVII (1989), 828-861.

________, "Debt Maturity Structure and Liquidity Risk," *Quarterly Journal of Economics*, CVI (1991), 709-738.

Dunn and Bradstreet, *Handbook of Credit Terms* (1970).

Elliehausen, Gregory, and John Wolken, "The Use of Trade Credit by Small Businesses" Board of Governors of the Federal Reserve (1992).

Fischer, Klaus, "Hausbankbeziehungen als Instrument der Bindungzwischen Banken und Unternehmen", Ph.D. Thesis, Universitat Bonn, (1990).

Gertler, Mark, "Financial Capacity and Output Fluctuations in an Economy with Multi-Period Financial Relationships," *Review of Economic Studies*, LIX (1992), 455-472.

Hannan, Timothy, "Bank Commercial Loan Markets and the Role of Market Structure: Evidence from Surveys of Commercial Lending," *Journal of Banking and Finance*, XV (1991), 133-149.

Hart, Oliver, and John Moore, "Default and Renegotiation: A Dynamic Model of Debt," Massachusetts Institute of Technology Working Paper, 1989.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein, "Bank Monitoring and Investment: Evidence from the Changing Structure of Japanese Corporate Banking Relationships," in R. Glenn Hubbard, ed., *Asymmetric Information, Corporate Finance and Investment*, (Chicago, IL: University of Chicago Press, 1990).

James, Christopher, "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics*, XIX (1987), 217-235.

James, Christopher, and Peggy Wier, "Borrowing Relationships, Intermediation, and the Cost of Issuing Public Securities," *Journal of Financial Economics*, XXVIII (1990), 149-171.

Kaplan, Steven, and Bernadette Alcamo Minton, "Outside Intervention in Japanese Companies: Its Determinants and Its Implications for Managers", University of Chicago, 1992.

Lummer, Scott, and John McConnell, "Further Evidence on the Bank Lending Process and the Capital Market Response to Bank Loan Agreements," *Journal of Financial Economics*, XXV (1989), 99-122.

Macrae, Norman, "Sweaty Brows Slippery Fingers," *The Economist*, (September 8, 1990), 21-24.

Meerschwam, David, *Breaking financial boundaries: Global capital, national de-regulation and financial services firms* (Cambridge, MA: Harvard Business School Press, 1991).

Mayer, Colin, "New Issues in Corporate Finance," *European Economic Review*, XXXII (1988), 1167-1189.*firms* (Cambridge, MA: Harvard Business School Press, 1991).

Mian, Shehzad, and Clifford Smith, "Accounts Receivable Management Policy: Theory and Evidence," *Journal of Finance*, XLVII (1992), 169-200.

Myers, Stewart, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, V (1977), 147-175.

Petersen, Mitchell, and Raghuram Rajan, "The Benefits of Lending Relationships: Evidence from Small Business Data," *Journal of Finance*, XLIX (1994), 3-37.

Rajan, Raghuram, "Insiders and Outsiders: The Choice between Informed and Arm's Length Debt," *Journal of Finance*, XLVII (1992), 1367-1400.

Stiglitz, Joseph, and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, LXXI (1981), 393-410.

Thakor, Anjan, and Richard Shockley, "Information Content of Commitments to Lend in the Future: Theory and Evidence on the Gains from Relationship Banking," Indiana University Discussion Paper, No 523, 1992.

Townsend, Robert, "Optimal Multiperiod Contracts and the Gain from Enduring Relationships," *Journal of Political Economy*, XC (1982), 1166-1186.

1. The basic argument dates back, at least, to Schumpeter who proposed that a monopolistic economy offers better incentives for innovation because an inventor can recoup her investment in R&D through future rents. Mayer [1988] introduces this problem to the finance literature discussing the general inability of firms and creditors to commit to a mutually beneficial course of action. Fischer [1990] models the problem posed by competitive credit markets formally and suggests that a way for a firm to commit to sharing future rents with the bank is to give the latter an informational monopoly. Rajan [1992] argues that even the bank's information monopoly may be insufficient to bind the firm when competition comes from arm's length markets.

2. The Economist [November 13, 1993, p 84] describes the problem as follows:
"...banks remain unable to charge prices that reflect the high risks of lending to small companies. Start-ups are especially risky; not only do plenty of them fail, but those that succeed tend to exploit their success by refinancing their debt more cheaply.... One big British bank reckons that to make a profit today it would have to double interest margins on its loans, from an average of three percentage points over the base rate. That might provoke the government to intervene.

So banks are looking for other ways to boost returns from borrowers that succeed. Some, such as Midland, would like to take small equity stakes. Others talk of introducing a clause into loan agreements that would give the bank a one-off fee if a borrower wanted to refinance its debt. Customers are understandably unkeen."

3. Bad entrepreneurs are the incompetent, the lazy, and the dishonest in the population of potential entrepreneurs. The incompetent invest in the wrong projects and consequently waste the investment, the lazy do not put in effort and see their investment wither away, and the dishonest steal the money or extract excessive private benefits from the firm.

4. In a more detailed model, we could motivate debt as a way of extracting cash from entrepreneurs, as for example in Hart and Moore [1989], or the optimal contract in a costly state verification setting as in Diamond [1984].

5. See Myers [1977] and Hart and Moore [1989] for a discussion of the difficulties in writing such contingent contracts.

6. The first inequality in (A.5) limits the number of cases we have to examine and does not change the qualitative implications. In a more detailed model, M could be the outcome of a bargaining game.

7. Our model is functionally identical to Becker [1975]. Our competitive credit market is his general training. Our concentrated credit market is his specific training.

8. In other words, the firm borrows short term each period. Clearly, the firm could borrow at date 0 and spread repayments between date 1 and date 2. Since the discount rate is 0, this has no effect in the model and is omitted for simplicity. Of course, we allow for the fact that the good entrepreneur may not have enough cash to repay the loan at date 1.

9. Note that the bank has no incentive to forgive debt (so long as expected cash flows are enough to meet the contracted repayments) at date 1, since there is no moral hazard or adverse selection problem at date 1.

10. This is because the rent on the funds lent conditional on the risky project succeeding is $(M-1)[I_{1R} - (R_1 - D_1)]$ which using (A.1) and (A.2) is easily shown to exceed the rent on the funds lent for continuing the safe

project $(M-1)[I_{1S} -(S_1 - D_1)]$. This result holds much more generally. All we need for it to go through is that the bank receive a larger amount from projects which have higher NPV (net present value). This would happen if the firm expands and directs more business towards the bank, or if the bank uses its bargaining power to extract some of the additional NPV for itself.

11. The bank cannot commit to lending at date 1, else the bad entrepreneurs waste even more money. Furthermore, a loan commitment contracted at a high rate will not work as the firm can always decide not to take it down.

12. Firms involved in the agriculture, forestry, and fishing industry; finance and insurance underwriting; or real estate investment trusts were excluded from the survey. Firms were initially sent a series of work sheets which listed the financial information which was going to be collected by the questionnaire. The work sheets were followed up by a telephone interview.

13. We classify commercial banks, savings and loans associations, savings banks and credit unions as banks. Finance companies, insurance companies, brokerage or mutual fund companies, leasing companies and mortgage banks are classified as non bank financial institutions. We also have loans made by non-financial firms. The remaining loans consist of venture capitalist loans (surprisingly few), loans from government agencies and otherwise unclassified loans.

14. Hannan provides evidence that the market for bank commercial loans is local in nature -- specifying which market the firm is in helps explain the rate charged. The unit of describing markets in his study is the Metropolitan Statistical Area. In our study, the unit can either be the Metropolitan Statistical Area or the county.

15. We are assuming stationarity of the survival process so that we can indeed draw correct inferences about changes in borrowing and interest rates with age from a cross-sectional sample. This assumption is not required when we test differences across markets holding age constant.

16. When we consider only firms which have obtained institutional financing the differences seem to widen. The mean size of such firms in most competitive markets is almost twice the size of such firms in the most concentrated markets ($1,444,327 in competitive markets and $768,868 in concentrated markets) and the difference is significant at the 5 percent level. The medians are also different, $170,000 versus $149,000, and this difference is significant at the 10 percent level. The mean operating profit to asset ratio is higher in the competitive market (0.58 versus 0.42) while the mean sales growth is also somewhat higher (0.34 versus 0.30) despite the fact that the firms in the competitive credit market are larger.

17. Clearly, the annualized rate is not that high if firms are allowed to stretch repayments beyond the due date. Conversely, it may be much higher if firms miss the discount but pay before the due date. Since the stated terms in an industry may differ from actual industry practice, Petersen and Rajan (1994) construct empirical measures of the actual period for which firms can borrow if they do not avail of discounts. To estimate the potential stretch available to trade credit borrowers, they calculate the difference in the DPO between firms that regularly take the early payment discounts and those that do not. For each industry, they determine the median DPO for firms that take less than 10 percent of the discounts they are offered and the median DPO for firms that take more than 90 percent of the discounts offered. The difference between these two numbers is an estimate of how long firms that do not take discounts stretch out their payments. For the retail industry it is 8.9 days. Based on the standard terms, firms that do not take the discount are paying an additional 2 percent for 8.9 days of credit, which translates to an annualized interest rate of 123 percent.

18. Why do trade creditors lend when financial institutions are reluctant to lend? Trade creditors may have a better ability to dispose of the collateral, greater incentives related to the margins they make on the product they are selling, greater leverage over the firm because of their ability to cut off supplies, or greater information about the firm [see Mian and Smith 1992]. Trade credit is presumably very expensive because firms are not in the business of lending, though this merits further research.

19. As of June 1986, there were five remaining unit banking states: Colorado, Kansas, Montana, North Dakota, and Wyoming. Texas was a unit banking state till 1985.

20. Some states limit branch banking without having unit banking laws. If we expand our indicator variable to include states with limits on branching, the banking law coefficient rises from 9.4 to 12.9. The coefficient on the Herfindahl index, however, is unchanged.

21. While Table V suggests that the structure of the credit market does influence the availability of credit to firms, it does not suggest that the difference in availability across market structure is substantially different for young and old firms as our model would indicate. As we see from Table II, much of the difference in debt levels across markets is for firms four years or less in age. In a concentrated market, the median (mean) discount taken by such firms is 75 (60). In a competitive market, the median (mean) discount taken by such firms is 50 (52). So the difference in medians is 25 and the difference in means is 8. In comparison, consider firms older than 15 years. In a concentrated market, the median (mean) discount taken by such firms is 100 (78.5). In a competitive market, the median (mean) discount taken by such firms is 100 (78.3). So the difference in medians is 0 and the difference in means is 0.2. While we have increased the power of the test (and the economic difference) by taking more extreme groupings, the difference in availability across market structure for young and old firms is still not statistically significant.

22. The results indicate that as compared to firms in the most competitive market, firms in the most concentrated market pay lower interest rates when young and higher rates when old. One way to verify that this finding is not a result of our functional form is to examine the raw data. The evidence in Table VI confirms our intuition. Another way to check our results is to graph the residuals by age for both the most concentrated and the most competitive markets. When we do this we find no evidence of mispecification.

23. As might be expected given the similar distribution of industries across markets, industry adjustments do not affect these results. When we subtract the median value of sales growth for all firms having the same two digit S.I.C. code from the sales growth number and then find means and medians across markets, the differences (or the lack of them) are qualitatively similar. This is true for net profits and gross profits also. We adjust the assets by dividing them by the median in the industry. Again the results are qualitatively similar.

24. The evidence from the Japanese liberalizations seems to support this view. Hoshi, Kashyap and Scharfstein [1990] find that firms with relatively high growth and high Tobin's q reduced their bank ties to borrow from the markets. This suggests that the relatively high quality firms left the banks to borrow from the markets. Furthermore, the Economist reports that while banks were initially eager to lend to small firms to replace the business lost to the markets, they were now charging them an exorbitant rate "thus hitting the firms that borrowed at the wrong moment, and crimping their plans for new investment" [The Economist, October 27, 1990, p74].

25. That the sale of an equity stake to the lender helps a firm share surplus with it in times of distress has been put forward as one explanation of how the German universal banking system evolved. According to this

view, the current stakes of German banks in large German firms are not part of a carefully planned strategy, but a historic accident. The German banks acquired their stakes when they helped ailing German firms out in the 1920s and 1930s (The Economist, November 16, 1991, p102).

26. Governments do actively restrain competition in the financial sectors in the early stages of a country's growth, though these restraints tend to take on a life of their own. Macrae [1990] has an entertaining description of the operation of the old government-sponsored bank cartels. Meerschwam [1991] describes the product market segmentation in lending implemented by the Meiji Restoration in 1868. Subsequent to the attempts to introduce more competition during the U.S. Occupation after World War II, government guidance reasserted itself through the Ministry of Finance and the Bank of Japan. He argues that until the opening up of the bond markets, " specialized financial institutions, operating in well-defined product markets... limited financing alternatives for the various corporations. In addition, a lack of price differentiation among products reduced the incentive for participants to break these relationships."

27. This discussion presupposes that governments have a choice in deciding when to deregulate their financial systems. For some countries like Japan, the decision has been partly forced because large companies went offshore to tap Euromarkets [see Meerschwam 1991, p. 135]. In other countries, the wealthy vested interests set up by cartelization will successfully oppose liberalization.

Table I
Summary statistics for firms classified by age and credit market concentration

Firm characteristics	Young firms: Most competitive	Young firms: Middle market	Young firms: Most concentrated	Old Firms: Most competitive	Old Firms: Middle market	Old Firms: Most concentrated
Firm solicited by financial instititution during the year (1=yes)	0.32[10]	0.32	0.26	0.46[1]	0.39	0.31
	0.00	0.00	0.00	0.00	0.00	0.00
	(0.46)	(0.47)	(0.44)	(0.50)	(0.49)	(0.46)
	166	488	966	112	516	972
Firm size (in $1000s) Book value of assets	863	1074	569	1296	1692	1171
	102	108	103	151	207	155
	(3247)	(7188)	(1836)	(2722)	(5379)	(3062)
	179	517	1010	117	554	1027
Gross profits / Assets	2.78[1]	2.18	1.88	2.39[1]	1.95	1.79
	1.61[1]	1.24	1.06	1.36[1]	1.66	0.97
	(2.96)	(2.46)	(2.21)	(2.76)	(2.31)	(2.21)
	170	491	965	114	520	973
Operating profits / Assets	1.14	1.05	0.63	0.61	0.62	0.64
	0.17	0.18	0.15	0.23[10]	0.14	0.13
	(4.84)	(5.74)	(2.49)	(1.19)	(1.93)	(2.88)
	170	491	965	114	520	973
Sales growth (1986-1987)	0.30	0.32	0.31	0.15	0.15	0.14
	0.09	0.10	0.08	0.04	0.05	0.03
	(0.98)	(0.92)	(0.91)	(0.63)	(0.68)	(0.58)
	140	435	843	109	495	934
Number of instits. from which firm borrows	1.35	1.40	1.37	1.28[1]	1.36	1.43
	1.00	1.00	1.00	1.00	1.00	1.00
	(0.59)	(0.62)	(0.62)	(0.51)	(0.65)	(0.69)
	99	291	655	71	323	624
Longest business relation with financial institution (in years)	4.10[1]	4.66	4.75	16.37	16.42	17.42
	4.00	4.00	4.00	15.00	15.00	15.00
	(3.09)	(2.93)	(2.95)	(9.81)	(10.60)	(11.05)
	179	517	1010	117	554	1027

The S.B.A. survey classifies the Metropolitan Statistical Area or county where the firm is headquartered into three categories on the basis of the commercial bank deposit Herfindahl index of the area. The most competitive markets are those with a value of the Herfindahl index of less than 0.10. The most concentrated market are those with a Herfindahl index of more than 0.18. The middle market are those markets where the Herfindahl index is between 0.10 and 0.18. Young firms are ten years old (the median age) or less. Old firms are more than ten years old. We report the mean, the median, the standard deviation, and the number of observations in each cell.

For each variable we tested the equality of the means in the most concentrated and most competitive markets. We tested the equality of the medians in the most concentrated and most competitive markets using the Wilcoxon Rank-Sum test. Both tests were done separately for young and old firms. Differences in the means and medians is noted for significance levels of 10 percent ([10]), 5 percent ([5]), and 1 percent ([1]).

Table II
Institutional Indebtedness of Small Firms

Firm age (in years)	Most competitive market: Firms with instit. debt (Percent / number)	Most competitive market: Debt/assets (Mean / std dev)	Middle market: Firms with instit. debt (Percent / number)	Middle market: Debt/assets (Mean / std dev)	Most concentrated market: Firms with instit. debt (Percent / number)	Most concentrated market: Debt/assets (Mean / std dev)
≤ 2	0.47 12	0.49 (0.36)	0.49 55	0.44 (0.46)	0.65 147	0.46 (0.38)
3-4	0.49 20	0.44 (0.35)	0.60 60	0.39 (0.49)	0.65 141	0.43 (0.30)
5-6	0.65 17	0.28 (0.26)	0.55 64	0.43 (0.56)	0.68 131	0.39 (0.36)
7-8	0.69 20	0.19 (0.18)	0.63 50	0.38 (0.40)	0.64 125	0.43 (0.43)
9-10	0.57 17	0.32 (0.45)	0.60 65	0.34 (0.32)	0.65 117	0.40 (0.34)
11-13	0.70 19	0.32 (0.22)	0.57 60	0.45 (0.37)	0.66 142	0.38 (0.58)
14-17	0.59 13	0.44 (0.64)	0.64 75	0.38 (0.48)	0.62 146	0.33 (0.28)
18-22	0.75 15	0.38 (0.30)	0.61 75	0.27 (0.22)	0.67 126	0.34 (0.28)
23-31	0.63 17	0.96 (0.84)	0.55 62	0.29 (0.24)	0.61 123	0.33 (0.27)
Greater than 32	0.33 7	0.24 (0.14)	0.57 54	0.30 (0.28)	0.47 87	0.25 (0.24)

Firms are classified by the concentration of the credit markets where the firm is located and by the age of the firm. Firm age is years since acquired by the present owners. The SBA survey classifies the Metropolitan Statistical Area or county where the firm is headquartered into three categories on the basis of the commercial bank deposit Herfindahl index of the area. The most competitive market has a Herfindahl index of less than 0.10, and the most concentrated market has a Herfindahl index of more than 0.18. The debt to asset ratio includes only debt from institutions in the numerator. Loans from owners and family are not included.

Table III
Trade Credit Usage Across Credit Markets with Differing Degrees of Competition

	Most competitive market Median (Number of Obs.)	Middle market Median (Number of Obs.)	Most concentrated market Median (Number of Obs.)
Fraction of purchases offered on credit.	0.90 (231)	0.90 (818)	0.90 (1597)
Days Payable Outstanding.	15.83 (160)	16.22 (594)	14.58 (1232)
Fraction of firms taking less than 10 percent of the discounts	0.33	0.29	0.19
Fraction of firms taking more than 90 percent of their discounts	0.50	0.52	0.59

The most competitive market has a Herfindahl index of less than 0.10, and the most concentrated market has a Herfindahl index of more than 0.18. Days Payable Outstanding is 365 * Accounts Payable / Cost of Goods Sold.

Table IV
Credit Availability and the Role of Credit Market Competition

Independent variable	(I)	(II)	(III)	(IV)
<u>Firm characteristics</u>				
Log(book value of assets)	7.21[1]	7.26[1]	7.10[1]	7.13[1]
	(1.70)	(1.70)	(1.79)	(1.79)
Operating Profits / Book Assets	5.88[5]	5.95[5]	5.22[10]	5.35[10]
	(2.79)	(2.80)	(2.82)	(2.84)
Firm is a corporation (0,1)	-8.96	-9.10	-10.84[10]	-10.54[10]
	(5.88)	(5.89)	(6.24)	(6.24)
Log firm age (in years)	8.69[1]	8.74[1]	10.19[1]	10.39[1]
	(3.35)	(3.35)	(3.73)	(3.73)
Firm is located in an MSA	-7.44	-6.93	-4.32	-4.97
	(5.45)	(5.47)	(5.75)	(5.78)
Sales growth (1986-1987)			4.40	4.41
			(3.86)	(3.86)
Firm is located in a unit banking state		9.36		
		(9.94)		
<u>Relationship characteristics</u>				
Log(length of longest relationship) (in years)	18.64[1]	18.42[1]	19.70[1]	19.64[1]
	(3.84)	(3.85)	(4.05)	(4.05)
Debt from financial service provider (percent)	2.54	2.11	3.37	3.51
	(5.74)	(5.75)	(6.03)	(6.03)
Number of institutions from which firm borrows	-11.62[1]	-11.56[1]	-11.40[1]	-11.35[1]
	(3.16)	(3.16)	(3.33)	(3.33)
<u>Financial market concentration</u>				
Herfindahl > 0.18 (0,1)	16.54[1]	17.21[1]	16.87[1]	19.05[1]
	(5.66)	(5.70)	(5.98)	(6.30)
Herfindahl < 0.10 (0,1)				11.48
				(10.35)
Number of observations	1459	1459	1339	1339
-Log likelihood	4120.1	4119.7	3743.0	3742.4

The dependent variable is the percentage of offered early payment discounts taken by the firm. The coefficient estimates are from a tobit regression with two sided censoring. 15 percent of the observations are censored at 0 percent; 45 percent of the observations are censored at 100 percent. Standard errors are in parenthesis. The regression also includes seven industry dummies, three region dummies, and a constant.

[1] Coefficient is significant at the 1 percent level.
[5] Coefficient is significant at the 5 percent level.
[10] Coefficient is significant at the 10 percent level.

Table V
Percent of early payment discounts taken classified by firm age and credit market competition

Firm characteristics	Most competitive	Middle market	Most concentrated
Firm age <= 10 years	53.4[5]	54.7	63.0
	50.0	60.0	90.0
	(44.6)	(44.6)	(41.5)
	93	290	593
Firm age > 10 years	68.4[10]	67.2	76.6
	100.0	95.0	100.0
	(41.7)	(41.5)	(35.9)
	69	331	640

We tested the equality of the mean percent of early payment discounts taken in the most concentrated and most competitive markets. The test is done separately for young (firms less than or equal to ten years old) and old firms (firms greater than ten years old). Differences in the means are noted for significance levels of 10 percent ([10]), 5 percent ([5]), and 1 percent ([1]).

Table VI
Borrowing Costs Across Market Structures

	Young firms (Age <= 10 years)	Old firms (Age > 10 years)
Most competitive market	11.50	10.64
(Herfindahl < 0.10)	(2.31)	(2.19)
	59	48
Most concentrated market	11.38	11.04
(Herfindahl > 0.18)	(2.34)	(2.36)
	477	403

The S.B.A. survey classifies the Metropolitan Statistical Area or county where the firm is headquartered into three categories on the basis of the commercial bank deposit Herfindahl index of the area. The most competitive markets are those with a value of the Herfindahl index of less than 0.10. The most concentrated market are those with a Herfindahl index of more than 0.18. Young firms are ten years old (the median age) or less. Old firms are more than ten years old. We report the mean interest rate, the standard deviation, and the number of observations in each cell.

Table VII
The Evolution of Borrowing Costs Across Market Structures

Independent Variable	I	II	III
<u>Credit Market Conditions</u>			
Herfindahl greater than 0.18 (0,1)	-1.285^{10}		-1.347^{10}
	(0.686)		(0.690)
Herfindahl between 0.10 and 0.18 (0,1)	-1.704^{5}		-1.759^{5}
	(0.721)		(0.724)
Log of firm age if Herfindahl is greater than 0.18	-0.154^{10}		-0.153^{10}
	(0.088)		(0.088)
Log of firm age if Herfindahl is between 0.10 and 0.18	-0.135		-0.134
	(0.127)		(0.127)
Log of firm age if Herfindahl is less than 0.10	-0.725^{1}		-0.732^{1}
	(0.273)		(0.273)
Log of firm age if firm is in an MSA		-0.240^{5}	
		(0.112)	
Log of firm age if firm is not in an MSA		-0.154^{10}	
		(0.090)	
<u>Interest rate variables</u>			
Floating rate loan (0,1)	-0.477^{5}	-0.459^{5}	-0.479^{5}
	(0.187)	(0.188)	(0.187)
Prime rate	0.278^{1}	0.272^{1}	0.277^{1}
	(0.032)	(0.032)	(0.032)
Term structure spread	0.011	0.014	0.010
	(0.084)	(0.085)	(0.084)
Default spread	0.357^{5}	0.364^{5}	0.352^{5}
	(0.153)	(0.153)	(0.153)
<u>Firm characteristics</u>			
Log(Book value of assets)	-0.295^{1}	-0.301^{1}	-0.295^{1}
	(0.048)	(0.048)	(0.048)
Total debt/book assets	-0.011	-0.024	-0.003
	(0.147)	(0.148)	(0.148)
Firm is a corporation	-0.158	-0.173	-0.159
	(0.145)	(0.145)	(0.145)
Number of banks from which firm borrows	0.379^{1}	0.378^{1}	0.377^{1}
	(0.088)	(0.088)	(0.088)
Firm is in an MSA	-0.021	0.074	-0.031
	(0.138)	(0.343)	(0.139)
Firm is in a unit banking state			-0.213
			(0.244)
Loan is from a bank (0,1)	0.100	0.140	0.107
	(0.193)	(0.192)	(0.193)
Loan is from a non-financial firm (0,1)	-1.148^{1}	-1.140^{1}	-1.144^{1}
	(0.367)	(0.367)	(0.367)
Number of observations	1277	1277	1277
Adjusted R^2	0.160	0.154	0.159

The rate quoted on the firm's most recent loan is the dependent variable. The regression also includes seven industry dummies, three regional dummies, six dummy variables for the type of assets with which the loan is collateralized, and an intercept. The prime rate is the rate at the time the loan was made. The term structure spread is the difference between the 10 year government bond yield and the 3 month t-bill yield at the time the loan was made. The default spread is the difference between the BAA corporate bond yield and the 10 year government bond yield at the time the loan was made.

[1] Coefficient is significant at the 1 percent level.
[5] Coefficient is significant at the 5 percent level.
[10] Coefficient is significant at the 10 percent level.

Figure 1

Firms are classified by their one digit SIC code. We graphed the percent of firms in each industry for each of the deposit market concentrations: most competitive (Herfindahl index less than 0.1), most concentrated (Herfindahl index greater than 0.18), and the middle marke

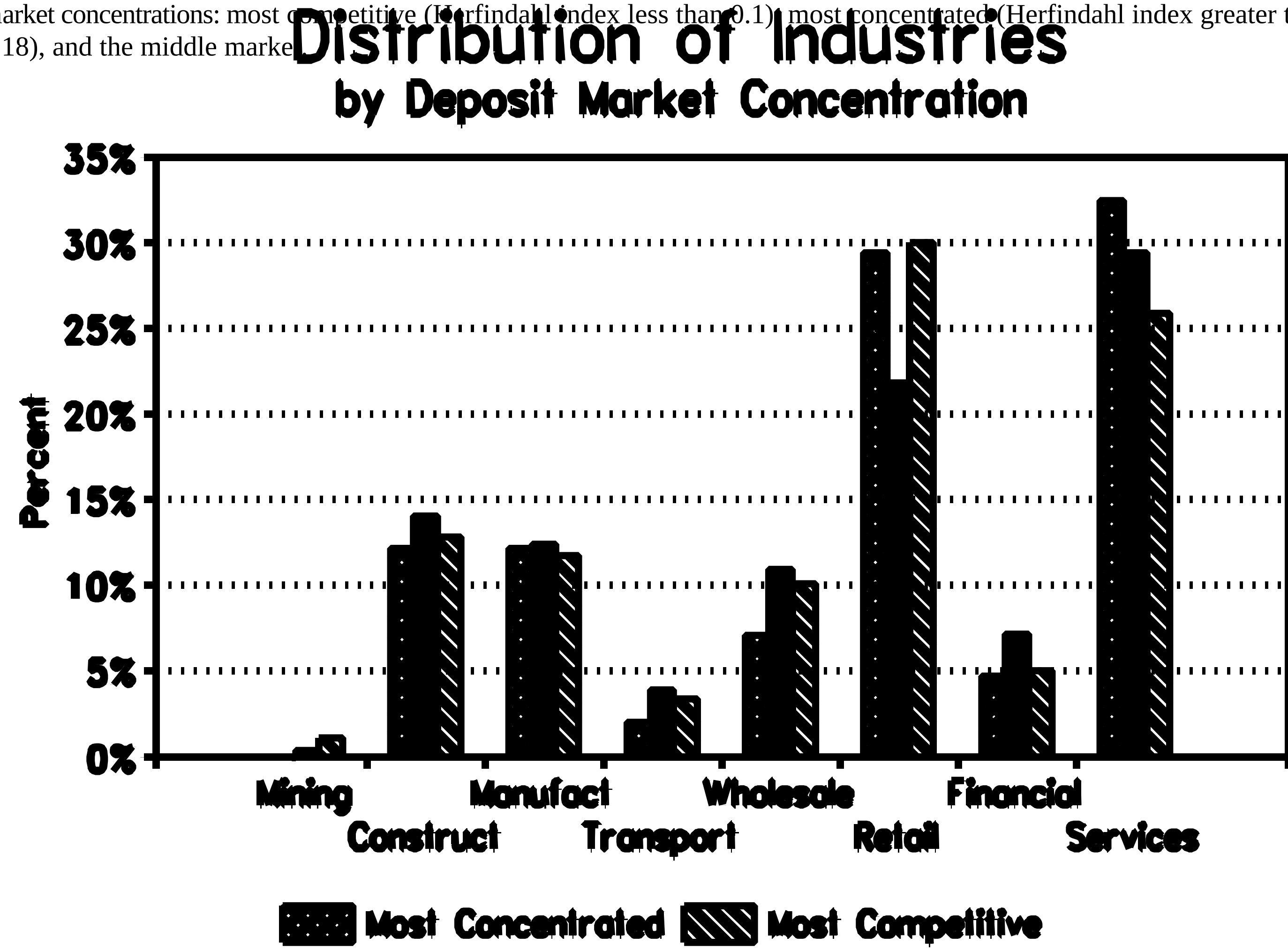
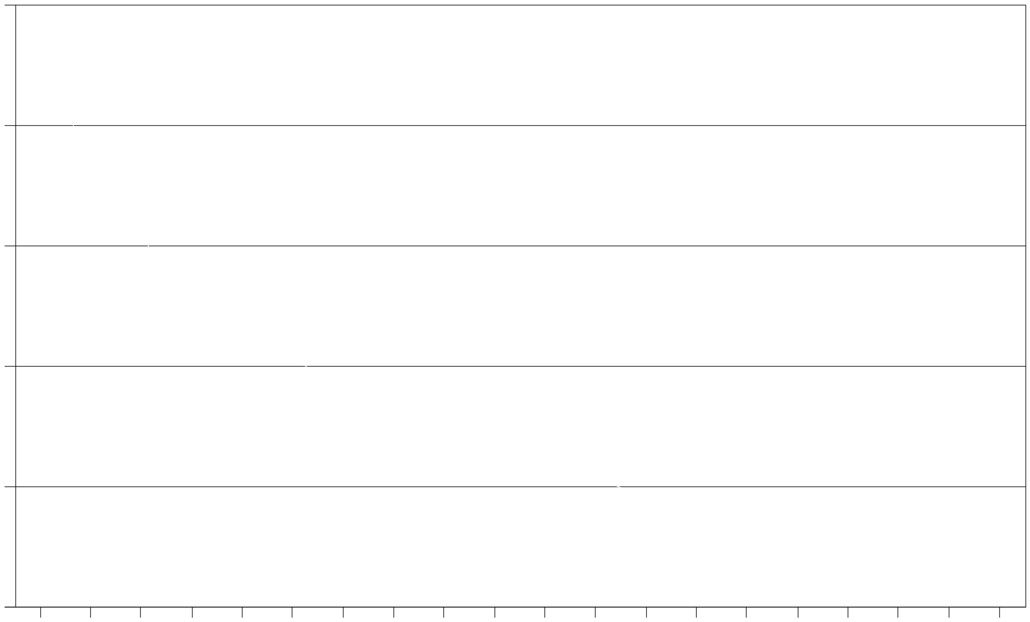


Figure 2

The loan rates are based on estimates from Table VII. All variables, except the firm's age and the market type (most competitive vs. most concentrated) were set equal to the sample means. The average loan rate in this sample is 11.1 percent.



Università degli Studi del Molise
Facoltà di Economia
Dipartimento di Scienze Economiche, Gestionali e Sociali
Via De Sanctis, I-86100 Campobasso (Italy)

ECONOMICS & STATISTICS DISCUSSION PAPER
No. 21/04

The Role of Guarantees in Bank Lending

by

Alberto Franco Pozzolo
University of Molise, Dept. SEGeS
and
Ente Luigi Einaudi

THE ROLE OF GUARANTEES IN BANK LENDING

By Alberto Franco Pozzolo[*]

Abstract

Guarantees play an important role in debt contracts. They alter the risk for the lender, transform borrowers' incentives and, possibly, modify the equilibrium allocation of financial resources. This paper studies the role of guarantees on bank loans, using a sample of over 50,000 individual lines of credit granted by Italian banks. Two empirical models are used. The first directly verifies the relationship between ex-ante publicly available information on borrowers' default riskiness and the presence of guarantees on their bank loans; the second compares the interest rates charged on secured and unsecured loans made by different banks to the same borrower, thus perfectly controlling for idiosyncratic riskiness and singling out the direct effect of the presence of guarantees on credit risk. The empirical results show that real guarantees (physical assets or equities that the lender can sell if the borrower defaults), which are often internal, are mainly used to provide a priority to some creditors. Personal guarantees (contractual obligations of third parties to make payments in case of default, e.g. suretyships), which can only be external, are used instead as incentive devices against moral hazard problems. Controlling for borrowers' characteristics, both real and personal guarantees reduce ex-ante credit risk.

JEL-classification: *G21, G32*

Keywords: Bank loans, collateral, guarantee

[*] Università degli Studi del Molise and Ente Luigi Einaudi.

1.Introduction[1]

A large number of bank loans are backed by collateral or guarantees.[2] Berger and Udell (1990) report that in the United States nearly 70 per cent of all commercial and industrial loans are made on a secured basis. Harhoff and Körting (1998) and Binks et al. (1988) report similar or even higher ratios for Germany and the United Kingdom, respectively.

The consequences of warranty requirements for the availability of bank financing have been examined in numerous theoretical and empirical studies. Information asymmetries in bank relationships can significantly alter the allocation of credit with respect to what would be socially optimal (i.e. all projects with a positive net present value – NPV – would be financed; see, e.g., de Meza and Webb, 1987). Warranties may help to alleviate these distortions by reducing moral hazard and adverse selection problems. They transform borrowers' incentives, alter the risks for the banks and eventually modify the equilibrium credit allocation. Smith and Warner (1979), for example, argue that "the issuance of secured debt lowers the total cost of borrowing by controlling the incentive for stockholders to take projects that reduce the value of the firm"; Stulz and Johnson (1985) show that in some cases recourse to

[1] I would like to thank Ugo Albertazzi, Allen Berger, Dario Focarelli, Andrea Generale, Giorgio Gobbi, Leonardo Gambacorta, Luigi Guiso, Simonetta Iannotti, Luigi Leva, Paolo Mistrulli, Fabio Panetta, Carmine Panzella, Bruno Parigi, Loriana Pellizzon, Salvatore Rossi, Stefano Siviero, Gregory Udell and seminar participants at the Bank of Italy, the University of Padua, the XIV Australasian Finance and Banking Conference and Ente Luigi Einaudi for their comments and suggestions, and Cinzia Chini and Stefania De Mitri for helping me through the data-bases. I alone am responsible for any remaining errors. I carried out most of this research while I was an economist at the Research Department of the Bank of Italy. Address for correspondence: Università del Molise, Dipartimento SEGES, via De Sanctis, 86100 Campobasso, Italy. E-Mail: pozzolo@unimol.it.

[2] There is no complete agreement in the literature on the definitions of guarantee and collateral. In the following I will use: a) guarantee for contractual obligations of a third party to make payments in case of default of the borrower (e.g. a suretyship), b) collateral for physical assets or a securities – posted either by the borrower or by a third party – which the lender can realize in case of default and c) warranty as a generic word indicating indifferently collateral and guarantees.

secured debt may make it possible to finance positive NPV projects that otherwise would not be financed.

However, warranties can also introduce new inefficiencies in credit allocation. For example, banks might devote less resources to screening and monitoring projects financed with secured loans, as the warranties themselves help to reduce credit risk (see, e.g., Manove et al., 2000). Consequently, if banks are more qualified than the average investor to evaluate projects, credit allocation may be less efficient when a larger proportion of loans is made on a secured basis. Moreover, if banks find it less expensive to require warranties than to monitor projects, investors who cannot provide them possibly may not be financed, even if the NPV of their investment is positive. Further, additional distortions might be introduced if some banks, observing warranty demands of other institutions, free-ride on their auditing activity. As shown by Rajan and Winton (1995), this may lead to sub-optimal monitoring.

One of the crucial issues in the analysis of secured bank lending is whether secured debt is requested at safer or riskier borrowers. This question has been answered in different ways in the light of the predictions of theoretical models, the conventional wisdom among bankers, and the results of econometric analyses. This paper provides some additional empirical evidence on the relationship between risk and warranties on bank loans, using high quality data on over 50,000 individual lines of credit granted by a large sample of Italian banks. It arrives at two main findings. First, borrowers with higher ex-ante probability of default are more likely to be required to post guarantees – which can only be offered by an external grantor – but not collateral – which is typically owned by the borrower. Second, controlling for borrowers' risk, secured loans carry lower rates than unsecured loans. This result is novel to the literature,[3] but it is consistent with the predictions of a large body of theoretical research and with the received view within the banking community.

The rest of the paper is organized as follows. The next section briefly summarizes the theoretical and empirical results of the literature on the relationship between borrowers' risk and secured bank loans. Section 3 discusses the hypotheses under scrutiny and the empirical models adopted. Section 4 describes the data used in the empirical analysis. Section 5 presents the results of the empirical analysis. Section 6 concludes.

2.Risk and warranties on bank loans

2.1 Theoretical results

The predictions of the theoretical literature on the relationship between risk and warranties strongly depend on the informational framework adopted.[4] Following the seminal contribution of Stiglitz and Weiss (1981), many models have been developed that assume that banks cannot observe borrowers' characteristics, so that the average interest rate on loans is higher than the rate that would be optimal for safe borrowers, if they could be identified. This creates an adverse selection problem, because only riskier borrowers apply for bank loans. In the original model, the equilibrium entails some degree of credit rationing. However, a possible alternative is to allow loan applicants to use warranties as a signaling device: by providing them, safer borrowers can credibly show their characteristics. Banks can therefore screen applicants by their degree of riskiness, offering better credit conditions to the safer ones. In this framework, secured loans are always those made to the safer borrowers, as shown by Bester (1985 and 1987), Chan and Kanatas (1985) and Besanko and Thakor (1987).

---

[3] One notable exception is Harhoff and Körting (1998). However, these authors do not expand in their finding.

The positive relationship between borrowers' riskiness and the presence of warranties on bank loans is a general result in models where they are used as a signaling device. Theoretical models where secured loans are made to riskier borrowers typically build on different assumptions. The most common, and probably the most compelling, is that warraranties are used as incentive devices in the presence of moral hazard problems. Boot et al. (1991) show that if the returns from the project that is financed depend, at least in part, on the degree of effort provided by the borrower – which is unobservable by the bank – and riskier applicants have a higher return from effort, then it is optimal for the bank to require a warranty from the riskier borrowers in order to limit moral hazard. Similarly, a moral hazard problem lies at the root of the results in Bester (1994), who shows that when the lender cannot credibly commit to forcing the bankruptcy of a borrower who cheats on the outcome of his investment, not repaying his debt, a warranty can be used to make the strategic default less attractive, therefore forcing the borrower to truly report his status. Because in equilibrium the incentives to strategically default are negatively correlated with project risk, banks will grant secured loans to riskier borrowers.

John et al. (2002) point to a different implication of the agency problems between managers and claimholders. Building on the seminal paper of Jensen and Meckling (1976), they show that if, in the event of default, the value of the assets posted as collateral is more stable than that of the other assets owned by the firm, managers have a stronger incentive to perk-consume secured than unsecured properties. As a result, equilibrium yields will be higher on collateralized than on uncollateralized debt, in order to compensate for the greater risk of "asset substitution".

Other authors have developed models where a positive relationship between borrowers' riskiness and the presence of warranties does not depend on moral hazard

---

[4] For a recent survey of the theoretical literature on the role of collateral in banking, see Coco (2000).

problems. Coco (1999), for example, shows that, even with ex-ante asymmetric information between borrowers and lenders as in Stiglitz and Weiss (1981), if borrowers are heterogeneous in their degree of risk aversion and those more risk averse are less willing to provide warranty, a screening equilibrium where guarantees are used as a signaling device is not possible and only risky borrowers may be requested to post collateral. de Meza and Southey (1996) show that when the population is composed of a number of overoptimistic borrowers, projects backed by high warranties are more likely to default. Finally, Barro (1976) shows that if the value of the warranty on bank loans is stochastic, and borrowers strategically default when its realization is lower than the sum of the value of the loan and its service, the equilibrium interest rate on secured loans is higher than that on unsecured loans, implying a positive correlation between risk and warranty. As suggested by Coco (1999), the same result can be explained by the presence of a ceiling on bank interest rates, for example due to usury laws.

2.2 Empirical evidence

The heterogeneity of results of the theoretical literature is shared only in part by the results of the empirical studies. Moreover, it is completely at odd with the conventional wisdom among bankers, who believe that banks typically require warranties on loans made to riskier borrowers.[5]

Some authors have checked whether secured loans have characteristics that plausibly signal them as riskier, considering a large number of variables.[6] The neatest result is that loans of longer duration are more likely to be secured, as found by Boot et al. (1991) and Harhoff and Körting (1998). With respect to the size of loans and

---

[5] See, for example, Morsman (1986).

[6] With a few exception (e.g., John et al., 2002), the theoretical literature does not distinguish between borrowers' and loans' riskiness. By contrast, the empirical literature takes account of loan-specific characteristics (e.g., size and duration), which are likely to affect credit risk for any given borrower.

borrowers, the results are less clear-cut. Harhoff and Körting (1998) and Elsas and Kranen (2000) find a higher incidence of securitization on larger loans – as one would expect considering that they typically entail a higher risk for the bank – but Boot et al. (1991) find a lower incidence. Beger and Udell (1995) find a positive relationship between the size of borrowing firms, measured by their total assets, and the probability that their lines of credit will be secured, and Harhoff and Körting (1998), proxying size with the firm's workforce, also find a positive relationship with the presence of warranties. On the other hand, the results of Elsas and Kranen (2000), showing a negative relationship between the presence of warranties and borrowers' total sales, are more in line with the conventional wisdom that smaller borrowers entail higher risk.[7] Harhoff and Körting (1998) also find that the share of secured loans decreases with the number of banking relationships, possibly because multi-banking wipes out the incentives to monitor borrowers' behavior or to require warranties of firms in financial distress, as suggested by Rajan and Winton (1995). Finally, Berger and Udell (1995) and Harhoff and Körting (1998) show that loans to borrowers with longer lending relationships – typically considered to be less risky – are less likely to be secured.[8] However, Elsas and Kranen (2000), using data from a survey of German banks, find that housebanks are more likely to have secured loans.[9]

Probably the most compelling results on the relationship between risk and warranties are those testing for differences in the interest rate spreads on secured versus unsecured loans. In a seminal contribution, using data from the FED survey on Terms of Bank Lending, Berger and Udell (1990) show that the interest rates on

---

[7] These differences may be due to the fact that the size of the borrower is related to his overall creditworthiness, which implies a negative relationship, but also reflects availability of assets to post as collateral, which implies a positive relationship.

[8] These results are consistent with the predictions of Boot and Thakor (1994), who show that an optimal contract implies that credit conditions become more favorable late in the relationship, after the borrower has already his ability to fulfill his obligations.

[9] Elsas and Khranen (2000) justify their result with the argument made by Welch (1997) and Longhofer and Santos (2000), who show that it is optimal for bank debt to be more senior when lending relationships are stronger.

secured loans are on average higher than those on unsecured loans. This result has two major implications: that secured loans are typically made to borrowers considered ex-ante riskier by banks, and that the presence of warranties is insufficient to offset such higher credit risk. Berger and Udell (1995) confirm this result using data on credit-lines from the same source. John et al. (2002), considering a sample of over 1,000 public issues of fixed-rate straight debt made between 1993 and 1995, find that yield on collateralized debt is higher than on general debt, even after controlling for credit ratings. Casolaro et al. (2002), studying a large sample of syndicated credit facilities between 1990 and 2001, also find that secured loans have larger interest rate spreads than unsecured loans.[10]

3.Hypotheses under scrutiny and empirical modeling

The theoretical literature yields straight testable hypotheses concerning the relationship between the riskiness of borrowers and loans and the presence of warranties. From the discussion above it is clear that when warranties are used in order to address adverse selection problems engendered by information asymmetries, their presence should have no relationship with the borrowers' default risk, because this information is not asymmetric. On the other hand, if warranties are used mainly to provide incentives to riskier borrowers in the face of moral hazard problems – and riskier applicants have a higher return from effort – their presence should be positively correlated with ex-ante measures of borrowers' default risk.

Two strictly related empirical models are used in order to test these hypotheses.[11] The first directly verifies the relationship between ex-ante publicly

---

[10] Harhoff and Körting (1998), on the contrary, using data from a survey of small and medium-sized German firms, find that the interest rates on secured loans are lower than those on secured loans.

[11] These models are derived from very simple theoretical assumptions. Disregarding agency problems, bank profits, π^B, are an increasing function of the interest rate, R, of the value of guarantees, C, and of the probability of repayment, p: $\pi^B = f(R, C, p)$ with $f'_R, f'_C, f'_p > 0$. Assuming zero profits in the banking sector, the previous expression implies: i) a positive relationship between the presence of collateral and loan riskiness (measured by one minus the probability of repayment), controlling for the

available information on borrowers' default risk and the presence of warranties on their bank loans, thus discriminating between adverse selection and moral hazard theories of secured lending. The second singles out the direct effect of the presence of warranties on credit risk, by comparing the interest rates on secured and unsecured loans made by different banks to the same borrower. Clearly, this measure is untouched by the indirect effect on interest rates originating from the differences in the characteristics of borrowers with secured and unsecured loans.

An important distinction to be made when testing the relationship between risk and collateralization is whether the collateral is owned by the borrower (inside collateral) or by an external grantor (outside collateral). As pointed out by Berger and Udell (2000), inside collateral simply reorders creditor priority in case of bankruptcy, giving secured lenders a specific claim on the pledged assets. By contrast, outside collateral is similar to an infusion of equity by the grantor, because it exposes him to the potential losses of the business.

The relationship between collateralization and borrower and loan riskiness differs depending on whether inside or outside collateral is provided. The theoretical literature shows that inside collateral is not a good signaling device, because it does not increase the potential loss for the borrower if he gets bankrupt. The only exception is if the pledge of his assets results in a welfare or profit loss for the borrower, for example because this limits his possibility to dispose of the assets in order to pursue new investment opportunities (as suggested by Smith and Warner, 1979) or for perk consumption (as suggested by John et al., 2000). A similar argument applies to collateral used as a tool to limit moral hazard on the part of borrowers. Therefore, any relationship between risk and the presence of collateral – , independent on whether positive or negative – can be expected to be stronger in the case of outside collateral.

---

interest rate; ii) a negative relationship between the level of the interest rate and the value of collateral, controlling for loan riskiness. Agency problems alter the previous relationship by introducing indirect effects, with the probability of default being made dependent on the presence of collateral: $p = g(C)$ with $g'_C > 0$ or $g'_C < 0$ depending on the mechanism at work.

Unfortunately, the distinction between internal and external collateral has not been considered adequately in the empirical literature, mainly because of the unavailability of data discriminating between the two types.

A further distinction, which is partly related to that between internal and external collateral, is between real and personal warranties. Real warranties (collateral) are typically physical assets or equities that the lender can sell if the borrower defaults, while personal warranties (guarantees) are contractual obligations of third parties to make payments if the borrower defaults (e.g., a suretyship). As pointed out by Berger and Udell (2000), guarantees typically operate like external collateral, only that they do not give control over specific assets but represent a generic claim on the entire wealth of the grantor, who thus has a large degree of freedom in using and possibly neglecting it.

The potentially different role of real and personal warranties depends on the outcome of two opposing forces. Collateral is potentially more powerful, because it is less easy to dispose of, but if it is inside, it does not increase the value of assets that the lender can withhold in case of default. Guarantees are less powerful because they can be more easily disposed of, but they are more powerful because they can only be external. An a priori ranking of these two effects is impossible; which one will dominate is therefore an empirical issue. The empirical analysis that follows does not discriminate between inside and outside guarantees, because this information is not available, but makes a distinction between real and personal warranties.

As anticipated above, the first model directly verifies the borrower and loan characteristics most often associated with secured lending, controlling for the interest rate on the loan. Two sets of control variables are included in the regression, describing the characteristics of the borrowers and of the lending relationship. The first set consists of a measure of each borrower's probability of default, other characteristics that might influence his riskiness (such as his share of physical over total assets and the number of his banking relationships) and proxies for the degree of information available on his creditworthiness (e.g. firm size and age). The second set consists of measures of loan-specific riskiness (such as its size) and the strength of the lending relationship (e.g. its duration). In order to control for characteristics specific to the lenders, dummy variables for each bank are introduced. In practice, the following discrete choice specification is assumed:

$$Pr\,(Y_{ij} = g) = f\,(X_{ij}, W_j, Z_i, K_j) \qquad g = 0, 1, 2 \qquad (1)$$

where: Y_{ij} equals 0 if the loan made by bank i to borrower j is unsecured, 1 if it is secured with real but not personal guarantees, 2 if it is secured with personal guarantees; X_{ij} is a vector of variables specific to the bank-borrower relationship; W_j is a vector of characteristics of the borrower; Z_i is a vector of bank-specific dummies; and K_j is a vector of dummy variables for the borrower's branch of economic activity and geographic location. The adoption of a discrete choice model is justified by the fact that the value of the collateral pledged on each loan is not significant information: except for very few cases, loans are either fully secured or unsecured.[12] Equation (1) is estimated using a multinomial logit specification.

The second model, inspired by Berger and Udell (1990), provides an indirect test of the relationship between riskiness and the presence of warranties. In particular, it verifies whether, controlling for borrower and loan riskiness, the interest rates on secured loans are systematically different from those on unsecured loans. Clearly, as discussed above, a negative relationship between interest rates and the presence of warranties is to be expected, because the lender's loss in case of default is reduced by the value of the collateral.

This second model is tested with a regression of each bank loan on two dummies, taking the value of 1 the loan is secured with real or personal guarantees, respectively:

$$i_{ij} = f(S_{ij}, X_{ij}, Z_i, W_j) \qquad (2)$$

[12] See, in particular, Section 4 and Table 1.

where i_{ij} is the interest rate on the loan made by bank i to borrower j; S_{ij} are two dummy variables taking the value of 1 if the loan is secured, respectively, with real and personal guarantees and 0 otherwise; X_{ij} is a vector of variables describing characteristics of the lending relationship; Z_i is a vector of bank-specific dummy variables; W_j is a vector of borrower-specific dummy variables.

4.Data and summary statistics

The empirical analysis uses information on lines of credit to a large sample of Italian non-financial firms. The data are taken from three sources: banks' supervisory reports to the Bank of Italy (*segnalazioni di vigilanza*), the Central Credit Register (Centrale dei Rischi) and the Company Accounts Data Service (Centrale dei Bilanci).[13] The first source is used for data on banks' balance sheets. The second contains information on single bank loans, the interest rates charged and the value of the assets pledged as warranties (distinguishing between real and personal); loans are recorded only when they are above a threshold level of ITL 150 million (around €75,000). The third source contains balance sheet information on a large number of non-financial enterprises. In particular, it includes a measure of the risk profile of the borrower – obtained, following Altman et al. (1994), as a numerical score from two discriminant functions – accessible to all banks affiliated with the Company Accounts Data Service. According to their score, firms are grouped into seven classes, from low risk (high security) to high risk (risk of bankruptcy).[14] Data for 1997 from the Central Credit Register have been used. In order to avoid simultaneity problems, lagged averages of the balance sheet information of the borrowers between 1993 and 1996 have been used.

[13] For a detailed description of the banks' supervisory reports to the Bank of Italy, the Central Credit Register and the Company Accounts Data Service see also Pagano et al. (1998).

[14] For a similar classification, see Sapienza (2003).

Tables 1-3 introduce the summary statistics for data from the sample of bank loans obtained by merging the information from the Central Credit Register and the Company Accounts Data Service. Table 1 presents some basic statistics by type of warranty. Loans secured with collateral are 2.1 per cent of all lines of credit; those secured with guarantees are 5.4 per cent.[15] The mode of the ratios of the value of the warranty to that of the loan is zero in all cases.[16] The value is 0 at the 95th percentile for collateral and 94.6 for guarantees; it is 99.4 per cent for collateral at the 99th percentile. These statistics show clearly that, when present, warranties normally cover the full amount of the loan. The requirement of warranties that cover only partially the value of the loan, which is largely suggested by the theoretical literature, seems to be irrelevant from an empirical point of view.[17]

Table 2 presents summary statistics on the ratio of the value of warranties to that of loans, with a breakdown by type of warranty, size of the lending bank, geographical area of activity and size of the borrower. The ratio of the overall value of collateral to that of loans is 5.2 per cent; it is 7.2 per cent for guarantees.[18] Larger banks make less recourse to collateral, and make a wider use of guarantees. Small borrowers have a larger share of loans covered by collateral, while the differences are smaller for guarantees. Finally, the share of secured loans shows a high variability across geographical areas.

[15] Unfortunately, information is only available on whether a guarantee is posted on a given banking relationship, but not on which loan is actually secured. In order to avoid attributing a guarantee to an unsecured line of credit made to a borrower that has another type of secured loan with the same bank (e.g., a term loan), banking relationships involving loans other than lines of credit are excluded from the sample used to construct Tables 1-3. On the contrary, they are included in the data used in the econometric analysis.

[16] For guarantees exceeding the value of the loan, the latter value has been used in the numerator.

[17] In fact, it is to be expected that when the warranty does not cover the full value it is either because the price of assets pledged has fallen in the meantime or because guarantees have also been provided. In the case of guarantees, for which this information is available, it is often found that their value exceeds that of the loan.

[18] These ratios are larger than those referring to the number of secured and unsecured loans, showing that larger loans are on average more likely to be secured.

Table 3 presents the breakdown by branch of economic activity of the borrower. The ratio of the value of collateral to that of loans ranges from 0.0 per cent for communication services to 10.5 per cent for rubber and plastic products. Corresponding ratios for guarantees ranges from 0.0 per cent for communication services to 16.0 per cent for construction.

5. Empirical results

5.1 Guarantees and ex-ante riskiness of borrowers

Table 4 reports the results of the estimates of the probability of loans being secured, distinguishing between the cases when only real guarantees are posted and when personal guarantees are present.

Estimates are performed on a sample of 52,359 loans; bank dummies and dummies for the area and the sector of activity of the borrower, included in the regression, are not reported in order to save space. The test for the independence of irrelevant alternatives (IIA), verifying that the multinomial logit framework is to be preferred to standard binomial logit regressions, is unable to reject the null hypothesis that the remaining alternative is irrelevant in the choice of whether or not to post collateral or a guarantee, respectively, on the loan.[19] The pseudo R-squared of the regression is 0.09.

In the case of collateral, the coefficient of the score on the borrower's probability of default is not significantly different from zero. This result is potentially consistent with models motivating the use of warranties with adverse selection problems, which imply no relationships between warranties and information available ex-ante to lenders, such as the score. By contrast, it is not consistent with the

[19] The test is an application of the Hausman specification test and verifies whether removing one option from the set of choices available (i.e., considering two separate logit regressions) systematically changes the parameter estimates; it is distributed as a χ^2 with as many degrees of freedom as parameters to be estimated (see Hausman and McFadden, 1984).

hypothesis that collateral is used as an incentive device in the presence of moral hazard problems, as in Boot et al. (1991).

Although the supposition cannot be directly verified, the absence of a significant relationship between ex-ante riskiness and the presence of collateral is probably due to the fact that collateral is mainly represented by assets internal to the borrowing firm. As such, it does not increase the loss suffered in case of default, and therefore has little effect on borrowers' incentives.

A justification for the use of collateral, consistent with the absence of a relationship with borrower's riskiness but not based on adverse selection problems, is that it provides a priority to some creditors; in case of default, a bank whose loan is secured with an internal warranty is more likely than other lenders to recover its assets.

The positive and significant coefficients of the length of the lending relationship and of the dummy variable for companies that are more than 20 years old are indeed consistent with the hypothesis that collateral is used to provide a priority. Moreover, they are consistent with the argument made by Longhofer and Santos (2000), that borrowers have an incentive to post collateral when lending relationships are stronger, because in this case banks are more inclined to help them in situations of financial distress.[20] Furthermore, one can expect that the need to put a specific creditor in a better position than others is likely to be lower when the borrower owns a large share of assets that can be withheld in the event of default. Consistent with this interpretation, the coefficient of the borrower's share of physical over total assets is negative and significantly different from zero in the case of collateral.

[20] Welch (1997) also suggests that because banks are better equipped to contest priority in financial distress, it is more efficient to give them higher seniority ex-ante. Extending this argument, one could say that banks with a stronger lending relationship are also in a better position than others to contest priority.

In the case of guarantees, the coefficient of the score on the borrowers' probability of default is positive, and significantly different from zero. This is consistent with the hypothesis that banks use guarantees as incentives, in the presence of moral hazard problems. On the other hand, it is not consistent with the hypothesis that banks use guarantees in order to address adverse selection problems.

Additional evidence of the fact that ex-ante riskier borrowers are more likely to be granted loans secured with guarantees is given by the negative and significant coefficients of the length of the lending relationship and of the dummy variable for companies that are more than 20 years old. In fact, older borrowers and those with a longer lending relationship are typically less risky, because they have a longer record – public and bank-specific – on which their expected performance can be judged.

The coefficients of the other control variables, with few exceptions, are consistent with the predictions of the theoretical literature.

Larger loans typically imply a higher credit risk for the bank. Table 4 shows that the coefficient of loan size is positive for both collateral and guarantees, confirming that real and personal warranties are used to reduce credit risk.

If warranties were used to give some creditors a better position in case of default, one would expect a positive coefficient of the number of banking relationships. In fact, the estimated coefficient is negative and significantly different from zero for both real and personal guarantees. As suggested by Rajan and Winton (1995), this apparently counterintuitive result is consistent with the hypothesis that banks are unwilling to require a warranty on their loans if this has the side effect of making the result of their screening activity implicitly available to competing lenders.

Loans to larger borrowers are more likely to be secured with guarantees and less likely to be secured with collateral. These results are likely to be the effect of opposing forces. On one side, a number of factors suggest that larger borrowers should be less likely to have secured loans. For example, they have more market power than smaller debtors when contracting loan conditions and they are normally less risky, because they are more subject to market's scrutiny and their balance sheets data are more easily available to outside observers. On the other side, smaller borrowers have a better ability to establish sound lending relationships, which often make it unnecessary to require warranties. Moreover, larger borrowers, which often belong to groups, are likely to have lower costs in using guarantees, because these are provided by the holding company or its subsidiaries.

Finally, the coefficient of the interest rate is positive in the case of collateral and negative for guarantees; in both cases it is significantly different from zero. The positive correlation between interest rates and the presence of collateral is probably due to a common factor, not adequately controlled for, driving both variables. A likely candidate is unobservable risk, coming from banks' private information about their borrowers' characteristics. In the following section this issue will be addressed in greater detail.[21]

5.2 Ex-ante riskiness of secured vs. unsecured loans

The results of the estimates verifying the borrower and loan characteristics most often associated with secured lending provide evidence in favor of the hypotheses that collateral is used primarily to provide a priority to some creditors over others, probably because it is largely internal, while guarantees, which are necessarily external, are used as an incentive device to reduce moral hazard problems.

The estimates of the model in equation (2), reported in Table 5, provide some additional evidence on the effect of warranties on loan riskiness. Clearly, a major problem in estimating the effect of warranties on a loan's interest rate is the potential

---

[21] Clearly, the presence of an uncontrolled common factor might also bias the estimates of the coefficients of the other explanatory variables (see Yatchew and Griliches, 1985, for a discussion of specification problems in discrete choice models). In particular, the bias should be positive for the coefficient of the borrower's probability of default, which is likely to be positively correlated with unobserved risk. As such, the effect of the score, the observable measure of borrowers' risk, would partly incorporate unobservable risk too. The bias should instead be negative for all the other coefficients, because loan's value, relationship length, company's age, number of banking relationship, borrower's total sales and borrower's share of physical to total assets are all likely to be negatively correlated with unobservable risk.

endogeneity; it is likely that riskier borrowers not only have higher interest rates but are required to back their loans with a warranty. If riskiness was not adequately controlled for (for example because banks have private information), this would lead to a positive relationship between interest rates and the presence of guarantees, a result in fact found by the vast majority of the empirical studies.[22]

In order to take care of the endogeneity problem, the regression presented in Table 5 controls for borrower and lender-specific characteristics, by introducing bank and firm-specific dummies, as well as for some characteristics specific to each lending relationship (its duration and the size of the loan). With this procedure, made possible by the large number of multiple relationships that distinguish the Italian banking system,[23] borrowers' characteristics – including unobservable risk – are perfectly controlled for.

The results show that, controlling better for borrower and loan riskiness than was possible in previous empirical studies, the presence of warranties reduces the interest rate on bank loans. In Table 5 both coefficients of the dummy variables for secured loans are negative and significantly different from zero.

This result is quite novel to the literature, but it is not unexpected: controlling for borrowers' risk, the first order effect of the presence of a warranty is to reduce the loss for the lender in case of default.[24] In fact theoretical explanations of a positive relationship between warranties and the interest rate charged build on the assumption that their presence has unobservable effects on loan riskiness. Controlling better for loan riskiness, it is therefore to be expected that the intuitive negative relationship between warranties and the interest rate is recovered.

Clearly, even when borrower-specific fixed effects are introduced, the control for loan riskiness is far from perfect, because the amount of private information on a

[22] As mentioned above, one notable exception is Harhoff and Körting (1998).

[23] On this issue see, in particular, Ongena and Smith (2000) and Detragiache et al. (2000).

borrower (and therefore his perceived riskiness) is idiosyncratic to each bank, hence to each lending relationship. However, the presence of unmeasured loan riskiness introduces a positive bias in the estimate of the coefficient of the dummies for the presence of warranties. In its absence, the negative coefficients reported in Table 5 should be even larger in absolute value.

6.Conclusions

The empirical evidence presented in this paper sheds some new light on the determinants of banks' secured lending, partly reconciling academic research and the conventional wisdom of practitioners.

Using unique data on lending relationships it has been possible: first, to discriminate between adverse selection and moral hazard theories of secured lending, by verifying the relationship between ex-ante publicly available information on borrowers' default riskiness and the presence of warranties on their bank loans; second, to single out the direct effect on credit risk of the presence of warranties, by comparing the interest rates charged on secured and unsecured loans made by different banks to a same borrower.

The evidence presented is consistent with the view that collateral and guarantees have a different role in loan contracts. Collateral is mainly internal and is therefore used essentially to provide a priority to some creditors with respect to others. On the other hand, it is less likely to be used as an incentive device in the presence of moral hazard problems, as the latter case would imply a positive relationship between their presence and borrowers' ex-ante riskiness, which is not found. Still, the presence of internal warranties reduces banks' credit risk, as is shown by the fact that, once other sources of riskiness are adequately controlled for, secured loans are charged lower interest rates.

---

[24] See the discussion in footnote 11.

Guarantees, which are mainly external, are typically used as incentive devices in the presence of moral hazard problems. They are more likely to be found in loans made to borrowers with an ex-ante higher probability of default. Like collateral, guarantees reduce credit risk, as is shown by the fact that secured loans are charged lower interest rates.

Table 1

Secured loans: summary statistics

(percentage values)

Collateral is typically represented by physical assets or equities; guarantees are contractual obligations of third parties to make payments in case of default of the borrower, such as a suretyship. Due to the absence of information on the type of loan that is secured with guarantees, banking relationships involving loans other than lines of credit are excluded. Source: Italy's Bank Credit Register, 1997.

	Collateral	Guarantees
Share of secured loans	2.1	5.4
Value at 95th percentile	0	94.6
Value at 99th percentile	99.4	100

Table 2

Value of warranties relative to that of total loans by duration, size of the lending bank, geographical area and type of warranty

(percentage values)

Ratio of the value of warranties to that of total loans in the class. For warranties exceeding the value of the loan, the latter value has been used in the numerator. For variable and sample definitions, see also the note to table 1. Source: Italy's Bank Credit Register, 1997.

	Collateral	Guarantees
Bank size		
Below 20^{th} percentile	8.3	6.7
Between 20^{th} and 40^{th} percentile	6.9	6.6
Between 40^{th} and 60^{th} percentile	5.5	7.4
Between 60^{th} and 80^{th} percentile	3.4	7.4
Above 80^{th} percentile	3.0	9.1
Borrower size		
Below 20^{th} percentile	9.2	7.7
Between 20^{th} and 40^{th} percentile	7.6	6.3
Between 40^{th} and 60^{th} percentile	6.7	6.1
Between 60^{th} and 80^{th} percentile	6.8	8.1
Above 80^{th} percentile	2.9	6.2
Area		
North-West	6.3	5.9
North-East	6.3	7.7
Center	2.7	7.8
South	6.7	9.7
Islands	4.6	6.1
Total	5.2	7.2

Table 3

Value of warranties relative to that of total loans by branch of economic activity of the borrower

(percentage values)

Ratio of the value of warranties to that of total loans in the class. For waranties exceeding the value of the loan, the latter value has been used in the numerator. For variable and sample definitions, see also the note to table 1. Source: Italy's Bank Credit Register, 1997.

Branch of activity	Collateral	Guarantees
Agriculture	2.8	13.4
Energy	0.6	1.8
Food and tobacco products	7.5	4.4
Textiles	8.2	3.9
Leather and footwear	5.6	10.3
Wood and furniture	6.9	2.2
Paper and publishing	3.4	7.4
Chemicals	3.8	2.5
Rubber and plastic products	10.5	5.7
Metallurgy	4.1	11.3
Metals	4.8	7.2
Machinery for ind. and agr.	6.4	8.6
Electrical machinery	2.0	5.0
Motor-cars and other transp. eq.	1.1	4.6
Other manufactures	7.5	1.2
Construction	8.7	16.0
Commerce	5.0	5.0
Hotels	6.5	8.6
Transports	3.7	7.9
Communication	0.0	0.0
Other services	5.2	8.0

Table 4

Determinants of secured lending

The dependent variable equals 0 if the loan is unsecured, 1 if it is secured with collateral and 2 if it is secured with guarantees (see footnote 1 and equation 1 in the text). Borrowers' total sales are four years averages between 1992 and 1996. Geographical, sector and bank dummies, not reported, are included in the regression. For variable and sample definitions, see also the note to table 1. The test for independence of irrelevant alternatives (IIA) is distributed as a chi-squared under the null hypothesis of no systematic differences between logit and multinomial logit estimate, with as many degrees of freedom as parameters to be estimated. *** indicates significance at 1 per cent level. ** at 5 per cent and * at 10 per cent.

VARIABLES	Collateral		Guarantees	
	Coef. *Std err.*	Sign.	Coef. *Std err.*	Sign.
Risk	-0.01		0.04	***
(index)	*0.02*		*0.01*	
Loan's value	0.72	***	0.15	***
(logs – millions of lire)	*0.03*		*0.01*	
Relationship length	0.17	***	-0.14	***
(log – years)	*0.04*		*0.03*	
Borrower's age	0.09	*	-0.38	***
(dummy variable)	*0.06*		*0.04*	
Number of banking relationships	-0.05	***	-0.01	***
	0.01		*0.00*	
Borrower's total sales	-0.47	***	0.11	***
(logs – millions of lire)	*0.03*		*0.02*	
Borrower's share of physical to	-0.70	***	-0.04	
total assets	*0.15*		*0.13*	
Loan interest rate	0.06	***	-0.02	***
	0.01		*0.01*	
Test of IIA (p-value)	0.00	1.00	4.78	0.99
No. of observations	52,359			
Pseudo R-squared	0.09			

Table 5

Warranties and interest rates on bank loans

The dependent variable is the level of the interest rate on the loan. Bank and borrower dummies, not reported, are included. For the definition of the variables and of the sample, see also the note to tables 1 and 4. *** indicates significance at 1 per cent level, ** at 5 per cent and * at 10 per cent.

VARIABLES	Coef. *Std. Err.*	Sign.
Collateral	-0.59	***
(dummy variable)	*0.09*	
Guarantees	-0.11	*
(dummy variable)	*0.08*	
Loan's value	-0.30	***
(logs – millions of lire)	*0.01*	
Relationship length	0.43	***
(log – years)	*0.03*	
No. of observations	67,829	
Adjusted R-squared	0.50	

References

Altman, E. I., M. Giancarlo and F. Varetto, 1994, “Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (the Italian Experience)”, *Journal of Banking and Finance*, 18, 505-529.

Barro, R. J. (1976), “The Loan Market, Collateral, and Rate of Interest”, *Journal of Money Credit and Banking*, 8, 839-856.

Berger, A. N. and G. F. Udell (1990), “Collateral, Loan Quality, and Bank Risk”, *Journal of Monetary Economics*, 25, 21-42.

Berger, A. N. and G. F. Udell (1995), “Relationship Lending and Lines of Credit in Small Firms Finance”, *Journal of Business*, 68, 351-381.

Berger, A. N. and G. F. Udell (2000), “Small Business and Debt Finance”, in Zoltan, J. A and D. B. Audretsch (eds.), *Handbook of Entrepreneurship,* forthcoming.

Besanko, D. and A. V. Thakor (1987), “Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets”, *International Economic Review*, 28, 671-690.

Bester, H. (1985), “Screening vs. Rationing in Credit Markets with Imperfect Information”, *American Economic Review*, 57, 850-855.

Bester, H. (1987), “The Role of Collateral in Credit Market with Imperfect Information”, *European Economic Review*, 31, 887-899.

Bester, H. (1994), “The Role of Collateral in Model of Debt Renegotiation”, *Journal of Money Credit and Banking*, 26, 72-86.

Binks, M., C. Ennew and G. Reed (1988), “Small Business and Banks: A Two Nation Perspective”, prepared for the Forum of Private Business, UK and the National Federation of independent Business, USA.

Boot, A. W. A. and A. V. Thakor (1994), “Moral Hazard and Secured Lending in an Infinitely Repeated Credit Game”, *International Economic Review*, 35, 899-920.

Boot, A. W. A., A. V. Thakor and G. F. Udell (1991), “Secured Lending and Default Risk: Equilibrium Analysis, Policy Implications and Empirical Results”, *Economic Journal*, 101, 458-472.

Casolaro, L., D. Focarelli and A. F. Pozzolo (2002), “Risk Retaining or Reputation: What Provides Certification on Bank Loans? Evidence from the Syndicated Credit Market”, mimeo, Banca d’Italia.

Chan, Y. and G. Kanatas (1985), “Asymmetric Valuations and the Role of Collateral in Loan agreements”, *Journal of Money Credit and Banking*, 17, 84-95.

Coco, G. (1999), “Collateral, Heterogeneity in Risk Attitude and Credit Market Equilibrium”, *European Economic Review*, 43, 559-574.

Coco, G. (2000), “On the Use of Collateral”, *Journal of Economic Surveys*, 2, 191-214.

de Meza, D. and D. C. Webb (1987), “Too Much Investment: A Problem of Asymmetric Information”, *Quarterly Journal of Economics*, 102, 281-292.

de Meza, D. and C. Southey, (1996), “The Borrowers Curse: Optimism, Finance and Entrepreneurship”, *Economic Journal*, 106, 375-386.

Detragiache, E., L. Guiso and P. Garella (2000), “Multiple versus Single Banking Relationships: Theory and Evidence”, *Journal of Finance*, 55, 1133-1161.

Elsas, R. and J. P. Krahnen (2000), “Collateral, Default Risk, and Relationship Lending: An Empirical Study on Financial Contracting”, CEPR Discussion Paper n. 2540.

Harhoff, D. and T. Körting (1998), “Lending Relationship in Germany – Empirical Evidence from Survey Data”, *Journal of Banking and Finance*, 22, 1317-1353.

Hausman, J. and D. Mc Fadden (1984), “A Specification Test for the Multinational Logit Model.” *Econometrica*, 52, 1219-1240.

Jensen, M. C. and W Meckling (1976) (1976), “Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure”, *Journal of Financial Economics*, 3, 305-360.

John K., A. W. Lynch and M. Puri (2002), “Credit Ratings, Collateral and Loan Characteristics: Implications for Yield”, *Journal of Business*, forthcoming.

Longhofer, S. D. and J. A. Santos (2000), “The Importance of Bank Seniority for Relationship Lending”, *Journal of Financial Intermediation*, 9, 57-89.

Manove, M., A. J. Padilla and M. Pagano (2000), “Collateral vs. Project Screening: A Model of Lazy Banks”, CEPR Discussion Paper no. 2439.

Morsman, E. Jr. (1986), “Commercial Loan Structuring”, *Journal of Commercial Bank Lending*, 68, 2-20.

Ongena, S. and D. C. Smith (2000), “What Determines the Number of Bank Relationships? Cross-Country Evidence”, *Journal of Financial Intermediation*, 9, 26-56.

Pagano, M., F. Panetta and L. Zingales (1998), “Why Do Companies Go Public? An Empirical Analysis”, *Journal of Finance*, 53, 27-64.

Rajan, R. (1992), “Insiders and Outsiders: The Choice between Informed and Arm's-length debt, *Journal of Finance*, 47, 1367-1400.

Rajan, R. and A. Winton (1995), “Covenants and Collateral as Incentives to Monitor”, *Journal of Finance*, 50, 1113-1146.

Sapienza, Paola (2003), "The Effects of Government Ownership on Bank Lending", *Journal of Financial Economics*, forthcoming.

Smith, C. W. and J. B. Warner (1979), "On Financial Contracting. An Analysis of Bond Covenants", *Journal of Financial Economics*, 7, 117-161.

Stiglitz, J. E. and A. Weiss (1981), "Credit Rationing in Markets with Imperfect Information", *American Economic Review*, 71, 393-410.

Stulz, R. M. and H. Johnson (1985), "An Analysis of Secured Debt", *Journal of Financial Economics*, 14, 501-522.

Welch, I. (1997), "Why is Bank Debt Senior? A Theory of Asymmetry and Claim Priority Based on Influence Costs", *Review of Financial Studies*, 10, 1203-1236.

Yatchew, A. and Griliches, G. (1985), "Specification Error in Probit Models", *The Review of Economics and Statistics*, 67, 134-139.

THE JOURNAL OF FINANCE • VOL. LXII, NO. 6 • DECEMBER 2007

How Laws and Institutions Shape Financial Contracts: The Case of Bank Loans

JUN QIAN and PHILIP E. STRAHAN*

ABSTRACT

Legal and institutional differences shape the ownership and terms of bank loans across the world. We show that under strong creditor protection, loans have more concentrated ownership, longer maturities, and lower interest rates. Moreover, the impact of creditor rights on loans depends on borrower characteristics such as the size and tangibility of assets. Foreign banks appear especially sensitive to the legal and institutional environment, with their ownership declining relative to domestic banks as creditor protection falls. Our multidimensional empirical model paints a more complete picture of how financial contracts respond to the legal and institutional environment than existing studies.

BANKS ARE THE DOMINANT SUPPLIERS of external finance in most economies across the world, with foreign lenders playing an increasingly important role (Demirgüç-Kunt and Levine (2001)). When lending to a company in an emerging economy, a bank must assess not only the credit quality of the borrower but also the risks due to weak laws or institutions. This paper estimates how laws and institutions affect the price and nonprice terms, as well as the ownership, of bank loans around the globe. Our results support the logic underlying the law and finance literature, pioneered by La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1997, 1998; LLSV hereafter), that some environments are more conducive to writing and enforcing financial contracts than others, and that better contracting leads to better outcomes.[1]

*Both authors are at the Carroll School of Management, Boston College and are affiliated with the Wharton Financial Institutions Center; Strahan is also affiliated with the NBER. We appreciate helpful comments from an anonymous referee and an associate editor, Franklin Allen, Thorsten Beck, Cliff Holderness, Ross Levine, Katharina Pistor, Andrei Shleifer, Amir Sufi, Bernard Yeung, and seminar participants at the Bank of England, Boston College, Bundesbank, Federal Deposit Insurance Corporation, the Federal Reserve Banks of New York and Philadelphia, Harvard, INSEAD, London Business School, Rice, Texas A&M, the American Finance Association meetings in Philadelphia, and the Financial Intermediation Research Society meetings in Capri, Italy. Geraldine Hough, Frank Corrado, Jean-Marc Donics, Mehrun Etebari, Daniel Yang, and Zhe Xu provided excellent research assistance. We gratefully acknowledge Florencio Lopez-de-Silanes and Richard Cantor for providing data on creditor rights and sovereign debt ratings, and Boston College for financial support. The authors are responsible for all remaining errors.

[1] LLSV, among others, document that stronger investor protection and more efficient institutions are correlated with better financial and economic outcomes. However, time-series evidence on financial development challenges some of the correlations documented by LLSV (e.g., Rajan and Zingales (2003)).

Specifically, we find that in countries with strong creditor protection, bank loans are associated with more concentrated ownership, longer maturities, and lower interest rates. In countries with weak creditor protection, our findings suggest that, maturity, a nonprice term, substitutes for interest rate (the pricing term) and controls borrower risk (Diamond (2004)). Thus, consistent with the law and finance view, strong creditor rights seem to enhance loan availability as lenders are more willing to provide credit on favorable terms.[2] Moreover, the impact of creditor rights on loans varies with borrower characteristics. For example, as creditor rights improve, loans are more likely to be secured by collateral, and this relation is stronger when firms have more tangible assets. Finally, we find that foreign banks appear to be especially sensitive to the legal and institutional environment, with their ownership declining relative to domestic banks as creditor protection falls.

Our research extends the emerging empirical literature on financial contracting by exploring in a multivariate setting how banks and borrowers set ownership and contract terms. Previous studies focus on a single dimension (e.g., the interest rate) of what is a very complex contract that depends not only on interest rates, maturity, collateral, and ownership, but also on a host of complex and heterogeneous covenants. While we cannot consider every aspect of such contracts, our multidimensional approach allows us to paint a more complete and nuanced picture of how the availability and terms of credit respond to the contracting environment.[3]

Using a sample of loans in 43 countries (excluding the United States), we estimate six reduced-form regressions, where the endogenous variables are loan contract terms and the exogenous variables represent country-level legal and institutional variables.[4] Specifically, we examine how the basic pricing term (the interest rate), two nonprice terms not yet examined in the literature (loan maturity and an indicator for whether the loan is secured), and three ownership variables (the number of lenders, the share of each loan held by domestic banks, and the share held by government-owned banks) vary with creditor rights and other country-level variables such as the rating on a country's sovereign debt, legal enforcement costs (as measured by legal formalism), and a country's level of financial and economic development. In our tests, we also control for firm (and loan) characteristics that are likely to shape contract terms through variation in

[2] This normative conclusion is consistent with earlier research based on variation in creditor protection across U.S. states stemming from differences in the housing exemption (e.g., Berkowitz and White (2004)), and is also consistent with current reforms of U.S. personal bankruptcy law.

[3] For example, Esty and Megginson (2003), Esty (2004), and Sufi (2007) consider only loan ownership in the U.S. and other countries, Claessens and Klapper (2005) examine frequencies of bankruptcies in 39 countries, and Giannetti (2003) links *private* firms' financing decisions (including bank loans) to institutional variables. Outside of lending, Kaplan, Martel, and Stromberg (2006) and Lerner and Schoar (2005) examine private equity contracts in developed and developing countries.

[4] Since most of the legal and institutional variables have little time variation (Djankov, McLiesh, and Shleifer (2007; DMS hereafter)), we interpret the results as reduced forms.

credit risk or loan demand, but the main focus is on the effects of country-level factors, and in particular, on creditor rights.[5]

We draw three broad conclusions from our findings. First, consistent with LLSV, strong creditor rights expand loan availability because in the presence of better legal protection during bankruptcy and reorganization, lenders are more willing to extend credit on favorable terms ex ante. For instance, where creditor protection is strong, bank loans tend to have longer maturities and lower interest rates.[6] In addition, where creditor rights are weak, loan ownership becomes more diffuse as lender risk falls with diffuse ownership both by increasing the extent of diversification and by increasing the ex post cost of restructuring, which reduces the incentive for borrowers to default strategically (Bolton and Scharfstein (1996)). We also find that creditor rights affect a loan's price, maturity, and collateral requirements more for firms with more hard assets (property, plant, and equipment), as an increased capacity to pledge assets makes collateral more effective, increasing creditor protection and in turn enhancing loan availability.

Second, creditor rights affect domestic banks (located in the same country as the borrower) and foreign banks differently. Foreign banks own a significantly smaller proportion of loans in countries in which creditor protection is weak. This relationship is most pronounced among larger firms, because foreign banks tend to be a more important supplier of credit for these types of firms; smaller firms, in contrast, tend to rely on domestic banks regardless of the legal regime. These results suggest that foreign banks rely more on formal legal mechanisms to restructure distressed debt, while domestic banks work out bad loans privately, consistent with other findings from emerging economies (e.g., Mian (2006)).

Third, loan ownership interacts with both the price and nonprice terms of loans. We find that decreased loan concentration (i.e., more banks participating in a loan) increases the likelihood that collateral is used. Diffuse loan ownership—which, again, may reduce strategic defaults—may therefore complement collateral as a mechanism to solve borrower control problems. Interest rates also tend to decline with the participation of domestic banks, suggesting their better ability to assess borrower risk and solve control problems than foreign banks. Esty (2004) finds similar results using a sample of project finance loans.

[5] In earlier versions of the paper, we analyze a larger sample of loans from 57 countries without using firm characteristics. Adding firm characteristics (obtained from Worldscope) reduces sample size, but the links between creditor rights and maturity, ownership concentration, and the ownership of foreign banks become stronger; only the statistical relationship between creditor rights and the likelihood of collateral is weaker in the smaller sample.

[6] Using a similar data set to ours, Bae and Ghoyal (2004) focus only on the relation between interest rates on loans and cross-country legal and institutional variables. They find no effect of creditor rights on loan rates. They also find a negative effect of property rights on the interest rates on loans, including loans made to Asian countries. We drop Asian loans in our interest rate regressions because rates on these loans are *not* based on a markup over LIBOR, as the loans from other regions are priced in Dealscan.

Our sample of bank loans comes from Dealscan, with information on the borrowing firm matched with and augmented by Worldscope. Given our sample of banks and firms, we acknowledge at the outset that the results in this paper reflect the effects of cross-country differences in laws and institutions on financial contracts between relatively large (and often foreign) banks and large borrowers.[7] Also, our data only allow us to explore the terms of loans *actually made*; we note that borrowers' access to credit may be limited where legal rights, property rights, corruption, or asymmetric information are most problematic (e.g., Stiglitz and Weiss (1981), Djankov, McLiesh, and Shleifer (2007; DMS hereafter)).

The rest of the paper is organized as follows. In Section I, we describe our key legal and institutional variables and discuss why they may affect the terms of bank loans. Section II describes our sample and reports the empirical results. Section III concludes the paper. Explanations of all the variables that we use in the paper are presented in the Appendix.

I. Determinants of Bank Loan Contracts: Legal Protection, Institutions, and Other Factors

A. Theoretical Framework

Within the incomplete contracts framework, theories of debt based on the transfer of control rights upon default (e.g., Aghion and Bolton (1992), Hart and Moore (1994, 1998)) imply that if creditors have more bargaining power, that is, greater ability to force repayment or take control of the firm in the event of default, they will extend credit on more favorable terms (lower interest rates, longer maturities) ex ante. For example, lenders can better control borrower risk if they know they will be able to seize collateralized assets, or credibly threaten to take these assets, ex post, in default. Diamond (2004) argues that debt maturity is an effective contracting tool in environments with weak legal protection and costly contract enforcement.[8] In addition, the optimal concentration of creditors reflects the trade-off of inefficient renegotiation following default against borrowers' incentive to default strategically (e.g., Bolton and Scharfstein (1996)).

[7] Loans made to small, private borrowers are likely to be influenced significantly by a set of social, cultural, and relationship variables (Petersen and Rajan (1994)) that are difficult to observe and control for across a large number of countries. Hence, individual country-level studies would offer a more useful way to understand financial contracting for small firms. For example, Allen, Qian, and Qian (2005) show that in China, despite poor legal protection and inadequate financing through standard channels, the private sector has been growing fast, relying on alternative governance mechanisms and financing channels such as those based on reputation and relationships.

[8] Using a sample of commercial property loans (that have similar loan ownership and contract terms as in our sample), Benmelech, Garmaise, and Moskowitz (2005) find evidence that tighter zoning restrictions, which lower the liquidity of collateral, reduce a property's debt capacity. Similarly, Benmelech (2006) finds evidence that debt maturity is reduced when collateral is less liquid using variation in track width from 19th century U.S. railroads.

In a cross-country setting, creditors' ability to force repayment and costs of enforcing contracts (through the legal system) can be measured with legal and institutional variables. In this regard, the law and finance literature has established links between legal and institutional variables and financial/economic outcomes. LLSV (1997, 1998) differentiate countries with legal systems derived from those in England, France, Germany, and Scandinavia. They find that countries with English common law and French civil law origins lie at the extremes: Countries with English origin (French origin) provide the strongest (weakest) legal protections to both shareholders and creditors (LLSV (1998, 2000)). Countries with English origin also appear to have better institutions, including less legal formalism and hence lower enforcement costs in the courts (Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2003; DLLS hereafter)), and less corrupt governments (LLSV (1999)). Better legal protection and institutions, in turn, seem to lead to better outcomes for the financial system.[9] In addition to these four legal origins, LLSV also characterize another group of countries as having "socialist origin"—the countries in this category comprise those that have evolved from socialist, central planning economies to market-based economies. Not surprisingly, countries with socialist origin are much closer to French-origin countries than to English-origin countries in terms of their legal and financial systems.

Despite the work of LLSV and others on the relations among legal origin, institutions, and financial/economic outcomes, much less is known about the sources of these relations.[10] In particular, does legal origin, through its influence on a country's laws and institutions, *cause* observed differences in outcomes? The study of financial contracts offers a potentially powerful means to answer this question, but so far we know very little about how legal and institutional variables affect specific contract terms. Our aim is to fill this gap in the literature by testing directly how differences in the contracting environment affect the terms in bank loans to large corporate borrowers. In essence, LLSV and DMS (2007) study the *extensive* margin, asking, for example, how the contracting environment affects the total supply of credit, whereas we study the *intensive* margin, asking how loan terms vary with the environment for those firms with access to credit.

[9] Relative to firms in French-origin countries, firms in English-origin countries have more dispersed shareholder ownership (La Porta, Lopez-de-Silanes, and Shleifer (1999)), rely more on external capital markets to raise funds (LLSV (1997)) and have higher values (LLSV (2002)). In addition to the law and finance literature, the finance and growth literature supports the view that financial system development stimulates economic growth (e.g., McKinnon (1973), King and Levine (1993), Jayaratne and Strahan (1996), Rajan and Zingales (1998)), while a third strand of literature attempts to establish the links among law, finance, and growth (e.g., Levine (1999), Demirgüç-Kunt and Maksimovic (1998)).

[10] Related to the LLSV results, there is evidence that legal origin plays a role in explaining differences in both financial systems and economic growth. For example, the English common law system is superior to the French civil law system in its ability to adapt to evolving economic conditions (e.g., Posner (1973), Beck, Demirgüç-Kunt, and Levine (2003)). As a result, firms in French civil law countries face higher obstacles in accessing external finance than those in English common law countries (e.g., Beck, Demirgüç-Kunt, and Levine (2005)).

It is important to point out that problems stemming from asymmetric information between borrowers and lenders are also important (supply-side) factors for loan contracts. In particular, higher interest rates, while an effective tool to "price" risk under symmetric information, may have adverse effects under asymmetric information by worsening the moral hazard problem of borrowers choosing riskier projects (Stiglitz and Weiss (1981)). Accordingly, mechanisms aimed at alleviating or eliminating information asymmetry would increase the willingness of lenders to extend credit (e.g., Jappelli and Pagano (1993)). It therefore makes sense to focus not only on interest rates but also on the non-price terms and ownership of loans.

Based on these ideas, we consider two sets of explanatory variables in our tests. The first set accounts for demand-side factors and credit risk factors by incorporating firm characteristics (industry indicators, firm size, profitability, leverage, rating, total amount of fixed assets, and multinational status) and loan characteristics (loan purpose, loan type, currency, etc.) in the regressions. These measures have all been employed previously in the banking literature and, for our purposes, should be thought of as control variables. For example, larger borrowers and borrowers with highly rated debt pay lower interest rates and are more likely to be able to borrow on an unsecured basis relative to smaller and less well-rated borrowers (Strahan (1999)).

The second set of variables consists of the country-level factors. These variables represent the supply-side factors that may determine the total amount of credit and the terms of credit that lenders are willing to provide to borrowers. We view legal protection of creditors as the most important determinant of the contracting environment's quality, and hence it is the main focus of the paper. We also include other country-level institutional variables such as a country's sovereign debt rating, information sharing mechanisms, survey-based measures of legal enforcement costs (as measured by legal formalism), financial and economic development, and legal origins.[11] In the next section we provide a description of each of these country-level variables and explain how we incorporate them into our empirical models.

B. Country-Level Explanatory Variables

B.1. Creditor Rights

The strength of creditor rights is of paramount importance for lenders in determining the degree of their exposure to borrower expropriation. Following LLSV (1998), we use a single index to measure a borrower country's overall level of creditor rights. The index begins at zero and adds one for each of the following conditions that holds: (1) Secured creditors gain possession of assets once the petition for reorganization receives approval (i.e., there is no automatic stay on creditors' ability to seize collateral); (2) secured creditors are ranked

[11] Country-level factors also include measures of a country's macroeconomic condition, including business cycles.

first in the distribution of proceeds; (3) there are restrictions such as creditors' consent for going into reorganization; and (4) (incumbent) management does not stay in control of the firm during the reorganization. A higher score indicates stronger creditor rights. It is important to point out that the creditor rights index is constructed based on whether the protection of creditors' rights (along one of the four dimensions) is provided by the borrower country's law, which is typically stable over time (DMS (2007)).

Most of our data are on syndicated loans, a market dominated by large multinational banks serving large borrowers. Many of these loans are arranged and negotiated in either London or New York, and the loan contracts often include a "choice of law" clause that allows the law of the United States or the U.K. to supersede the laws in the borrower's country. Thus, to the extent that attempts to export U.S. or U.K. law succeed in nondefault states, our results will tend to be attenuated. However, the choice of law clause governs the credit or loan contract; it does not extend to bankruptcy, and typically not to property. The law that governs property relations or bankruptcy is determined by the borrower country's legal codes (i.e., where the assets are located) or in the local bankruptcy codes or insolvency laws (i.e., where the case is filed).[12] Consistent with empirical evidence from Esty and Megginson (2003), we find strong evidence on the borrower countries' legal codes affecting the ownership and terms of loan contracts around the globe.

B.2. Other Institutional Variables

While we focus our analysis on the impact of creditor rights on loans, we also test how legal formalism and information sharing registries affect loan ownership and contract terms. As we have discussed before, these (country-level) institutional variables, along with creditor rights, are potentially important determinants of the contracting environment for lenders and borrowers.

Legal Formalism: Legal formalism measures how efficiently the courts of the borrower's country enforce contracts (DLLS (2003)). Court efficiency matters because the ability of lenders to enforce or to threaten to enforce specific clauses of a loan contract (e.g., covenants), or to seize collateral, depends on the costs of using the legal system. Based on extensive surveys of lawyers and judges (see Appendix A.2 for more details), DLLS construct measures on how courts handle two types of cases: collection of a bounced check and eviction of a (nonpaying) tenant. A higher score in either category implies that the court system is slower (more bureaucracy) and less efficient. Although these measures are highly positively correlated across countries, we use the check-based

[12] See, for example, Norton (1996) for more descriptions on the legal environment of international syndicated markets, and Westbrook (2000) for descriptions on the legal complexities involving crossboarder bankruptcies. Siegel (2005) and Bergman and Nicolaeivsky (2007) show how Mexican firms (through cross-listing in the U.S. and adjusting corporate charters) voluntarily follow U.S. securities laws and protect investors beyond what is required in Mexico.

formalism index because the process of collecting a check boils down to enforcement of a financial contract.[13]

Information Sharing: There are two types of information sharing mechanisms (Jappelli and Pagano (2002)). First, a *public* registry is owned by public authorities (usually the central bank or banking supervisory authority), which collect information on the standing of borrowers in the financial system and make the database available to financial institutions. Second, a *private* bureau is a private commercial firm or nonprofit organization that maintains a database on the standing of borrowers in the financial system and facilitates exchange of information among banks and financial institutions. In our empirical models, the information sharing variable is equal to one if either a public registry or a private bureau operates in a borrower country, and zero otherwise.

B.3. Legal Origin

As we have discussed above, there are other aspects of legal protection and institutional efficiency that we do not consider directly, but that can affect the design of loans. Since LLSV demonstrate that almost all of these variables are highly correlated with indicators of legal origin, we can also interpret these legal origin indicators as proxies for possible omitted variables.[14] Hence, we include indicator variables in all of our specifications to control for legal origin.

B.4. Country-Level "Outcomes" Variables

Our aim is to isolate whether variation in laws and institutions changes the contracting environment and in turn the supply of credit. Thus, we need to control for demand variation in our regressions, as well as the variation in the risks of lending to borrowers across different economies. To ensure that effects of our institutional variables really reflect supply-side factors, we run some specifications that control for economic and financial development (to absorb unmeasured variation in credit demand) and a country's sovereign debt rating (to absorb unmeasured variation in risk). However, Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004) argue that country-level variables such as financial and economic development (as well as survey-based variables such as property rights protection) may be "outcomes" of legal origin and creditor protection. Controlling for such outcomes might therefore attenuate the total impact of

[13] DLLS also show that greater legal formalism (a higher score in either index) is highly *negatively* correlated with other measures of judicial quality, including enforceability of contracts, law and order, duration of trial and enforcement, and the extent to which a legal system is fair, uncorrupt, consistent, and affordable. These results imply that the check-based formalism index proxies for the costs of enforcing contracts as well as the effectiveness of a country's judicial system.

[14] For example, Stulz and Williamson (2003) find that culture (proxied by religion) is an important determinant of creditor protection, while DMS (2007) show that the impact of religion on creditor rights is absorbed by legal origins.

laws and institutions on loan contract terms, so we run all of our regressions with and without these variables.

Sovereign Debt Ratings: Because rating agencies assess the risk of sovereign (government) bond default, these ratings serve as a broad proxy for overall country risk. We therefore construct an index on country risk using Moody's ratings on the long-term sovereign bonds for the borrower's country (denominated in dollars). Specifically, we assign an Aaa rating a value of one, an Aa rating a value of two, . . ., and a B or worse rating a value of six; thus a higher number indicates a *lower* rating. We also assign a value of zero for countries without a sovereign debt rating (ratings are missing for about 15% of our observations), and we include an additional indicator for those firms for which the rating is missing.

Financial and Economic Development: To measure financial development, we include the ratio of total private domestic bank credit to GDP, averaged over the 1994–1997 period. For economic development, we use the log of GDP per capita, averaged over the 1994–2002 period.[15]

II. Data and Results

A. Data

A.1. Sample

We build a sample of bank loans made to large borrowers located in 43 countries (excluding the United States). Loan information comes from the Loan Pricing Corporation's Dealscan database, which provides detailed coverage of bank lending to large corporations. The data set goes back to the late 1980s, but coverage of lending to companies outside the United States is quite sparse until the middle of the 1990s. Hence, we begin our sample with loans originated in 1994 and include loans originated through the middle of 2003.

Dealscan allows us to identify which banks are lending to which firms in each year, and to observe various terms of the loans at origination, including the interest rate (measured as a basis point spread over LIBOR, inclusive of all fees), whether or not the loan is secured, the maturity of the loan, the number of lenders involved (many of the loans are syndicated), as well as the names

[15] Motivated by prior work (e.g., Acemoglu and Johnson (2005)), we include additional survey-based institutional variables measuring property rights protection and government corruption in an earlier draft of the paper. These variables are problematic for two reasons. First, like the sovereign debt rating or GDP per capita, they measure outcomes, so their effects combine both loan demand and supply factors. Second, including too many country-level variables induces multicollinearity and model instability because the country effects do not vary across loans to borrowers in the same country. Hence, it is difficult to sort out statistically the effects of all of the institutions on loan contracts. The effects of legal formalism, for example, are particularly sensitive to including these outcome variables. The impact of creditor rights on loan ownership and maturity, however, is robust to the inclusion of these variables and including these variables increases the significance of creditor rights in the secured lending regressions and decreases its significance in the interest rate regressions.

of the borrower and lender(s). From these last two pieces of information, we construct the share of the loan held (at origination) by domestic banks (located in the same country as the borrower), and the share held by banks controlled by governments (although not necessarily the borrower's government).[16] These ownership and contract features serve as the basis for the dependent variables in our models. To capture secured lending, we include an indicator equal to one if the loan is secured and zero if it is not.[17] For this qualitative variable, we estimate a probit and report *marginal* effects (rather than probit coefficients) in the tables. Thus, the reported coefficients represent the change in the probability per unit change in the relevant explanatory variables. (For indicator variables, the coefficient represents the change in the probability associated with moving the indicator from zero to one.) The other models are estimated with ordinary least squares (OLS), where the dependent variable equals the log of the number of lenders, the percentage of the loan held by government banks (ranging from zero to 100), the percentage held by domestic banks (zero to 100), the log of the loan maturity (in months), and the log of the basis point spread of the loan interest rate over LIBOR.

Beyond these loan contract terms, Dealscan also includes information on the type of loan (e.g., line of credit versus term loan), the purpose of the loan (e.g., debt repayment, commercial paper backup line of credit, general corporate purpose, etc.), and the size of the loan.[18] We control for loan type and loan purpose with indicator variables. We also include an indicator variable equal to one for loans denominated in one of the following major currencies: U.S. dollars, euros, German DM, British pounds, or Japanese yen. In addition, we include a regional indicator for firms located in Africa.

With respect to borrower characteristics, Dealscan contains information on Moody's and S&P senior debt ratings at the close of the loan, which we use to control for borrower risk. We focus on Moody's rating unless it is missing, in which case we rely on the S&P rating. Similar to sovereign ratings, we index firm ratings from one to six, with one indicating an Aaa rating, two indicating an Aa rating, . . . , and six indicating a B or worse rating; thus, again, a higher number indicates a lower rating. We assign a zero for borrowers without a rating, and we include a separate indicator for unrated firms. We include an indicator for multinational firms (a company is classified as a multinational if

[16] Information is based on the Bankscope database. Banks are considered government controlled if more than 25% of the shares are owned by governments. The home country of a bank is based on the location of the bank holding company. For example, Citibank's affiliate operating in India is classified as a U.S. bank.

[17] Due to data limitations, we do not have sufficient information on the characteristics of the collateral. Davydenko and Franks (2005) find that French banks respond to a creditor-unfriendly bankruptcy code by requiring more collateral (in value) than lenders in the U.K. and Germany, and by relying on particular collateral forms that minimize the statutory dilution of their claims in bankruptcy. Despite more collateral for small firms, however, recovery rates per dollar of debt were lowest among French defaults.

[18] The Dealscan data allow us to observe the commitment amount at loan origination, but not the amount of funds actually taken down by the borrower.

its parent and the company in question operate in different countries), a set of two-digit SIC indicator variables, and a set of year indicator variables. We drop loans to firms in SIC 6 (financials) and 9 (public sector) because firms in these sectors are more apt to be government-owned or government-protected monopolies, and the risks of these firms may be very different from those of manufacturing and trade firms.

Given the banks and firms covered in the Dealscan database, we acknowledge that our results will reflect the effects of cross-country differences in law and institutions on financial contracting between large banks and large borrowers. To some extent, focusing on these data limits the generality of our findings, but loans to smaller and privately owned firms are likely to be influenced by social, cultural, and relationship variables that are difficult or impossible to observe and control for across a large number of countries (Petersen and Rajan (1994)). In our view, individual country-level studies, rather than cross-country comparisons, would offer a more fruitful way to understand financial contracting for small and private firms.

To augment the Dealscan data on borrower risk, we hand-match our loan-level data to firm-level data from Worldscope based on the name of the borrowing firm. Worldscope includes balance sheet and income statement information for large, publicly traded firms across a wide range of countries. We use these financial data to construct measures of firm size (the log of the market value of equity in dollars), profitability (net income divided by assets), leverage (total debt divided by assets), and asset tangibility (property, plant, and equipment, or PP&E, divided by assets). We are able to match about half of the Dealscan firms to Worldscope. The borrower financials are measured at the end of the year in which the loan was originated, although this timing choice has little impact on the results (see the robustness tests below).

A.2. Summary Statistics

Table I reports median values for the loan terms for each of the 43 borrower countries in our sample as well as the key country-level variable (creditor rights). We group the medians by legal origin. The median loan size varies significantly across borrower countries, from \$25 million to \$40 million in developing countries such as China, India, and Pakistan, to \$355 million in Denmark. Much of the difference in loan size reflects difference in borrower size, as can be seen by comparing the median loan size with the median market value of equity (the simple correlation equals 0.64). Loan pricing also varies considerably across countries, ranging from a median spread of around 50 basis points in Hong Kong and Israel (both of English legal origin) to over 225 basis points in Brazil, Colombia, and Peru (all of French legal origin). These large pricing differences likely reflect, at least in part, the differences in average borrower risk for this latter group (the correlation between median spreads and borrower size equals –0.38.)

For median nonprice terms, the share of loans that are secured ranges from lows of 0% in Pakistan and 2.5% in Japan, to highs of 100% in a number of

Table I

Summary Statistics for Loan-Level and Country-Level Variables

This table reports country-level medians on loan terms by borrower country, grouped by legal origin (English, French, German, Scandinavian, and Socialist). The data sources are LPC's Dealscan and Worldscope. The sample includes loans originated between 1994 and the middle of 2003. We drop loans to financial firms, as well as public sector loans. The scores on creditor rights indexes are presented in the last column. Detailed descriptions and sources for the variables are presented in Appendix A.1.

Country	Number of Loans	Number of Lenders	Government Ownership (————	Domestic Ownership Percentage	Loan Secured ————)	Loan Maturity (Months)	Loan All-in Spread (Bps)	Loan Size (Millions	Market Capitalization of Dollars)	Creditor Rights
ENGLISH										
Australia	334	4	0.0	25.0	11.5	60	–	$135	$1,572	3.0
Hong Kong	157	1	0.0	0.0	14.5	48	50	64	423	4.0
India	98	3	0.0	4.9	35.5	60	–	40	541	2.0
Ireland	12	12	0.0	9.1	–	48	143	203	396	1.0
Israel	6	6	29.6	17.7	50.0	60	55	75	1,334	3.0
Malaysia	254	2	0.0	33.3	22.5	60	–	72	579	3.0
New Zealand	75	3	0.0	0.0	7.9	48	–	77	674	4.0
Pakistan	17	2	0.0	0.0	0.0	24	–	25	82	1.0
Singapore	117	1	0.0	11.1	20.8	54	–	46	313	3.0
South Africa	31	11	0.0	12.9	100.0	48	115	115	1,973	3.0
Thailand	201	3	0.0	0.0	16.1	60	–	46	286	3.0
U.K.	662	6	0.0	21.1	37.9	60	65	201	1,093	4.0
*Group Mean**	*164*	*4*	*2.5*	*11.3*	*28.8*	*53*	*86*	*92*	*772*	*2.8*
FRENCH										
Argentina	46	6	0.0	0.0	20.0	36	150	100	3,490	1.0
Belgium	11	8	1.7	22.2	100.0	59	90	152	2,039	2.0
Brazil	83	5	0.0	0.0	83.3	24	225	100	755	1.0
Chile	64	4	0.0	0.0	0.0	60	59	158	2,948	2.0
Colombia	17	6	0.0	0.0	33.3	60	275	88	1,109	0.0
France	146	8	0.0	50.0	41.2	59	80	185	4,044	0.0
Greece	22	8	0.0	41.5	100.0	54	140	80	253	1.0

Indonesia	35	11	0.0	0.0	40.0	36	–	50	134	3.0
Italy	39	4	0.0	14.6	60.0	36	40	173	2,836	2.0
Mexico	124	7	0.0	0.0	47.1	60	200	100	1,195	0.0
Netherlands	154	8	0.0	20.0	60.9	60	93	229	1,416	3.0
Peru	10	9	0.0	5.1	100.0	48	250	118	409	0.0
Philippines	140	3	0.0	0.0	20.0	60	–	38	1,273	1.0
Portugal	11	17	5.4	14.3	100.0	60	30	311	6,505	1.0
Spain	102	2	0.0	37.2	50.0	60	78	145	965	2.0
Turkey	13	7	0.0	0.0	–	60	150	75	832	2.0
Venezuela	5	10	10.0	10.0	100.0	51	238	113	784	3.0
*Group Mean**	*60*	*7*	*1.0*	*12.6*	*59.7*	*52*	*140*	*130*	*1,823*	*1.4*
GERMAN										
Austria	5	2	0.0	0.0	–	54	225	160	928	3.0
Germany	106	14	8.4	44.1	57.1	60	60	271	1,255	3.0
Japan	710	5	0.0	100.0	2.5	12	–	82	1,055	1.0
Korea (South)	301	2	0.0	26.7	7.9	36	–	50	772	3.0
Switzerland	40	14	8.3	11.7	20.0	24	54	245	1,849	1.0
Taiwan	10	3	1.6	100.0	44.4	60	–	103	1,798	2.0
*Group Mean**	*195*	*7*	*3.1*	*47.1*	*26.4*	*41*	*113*	*152*	*1,276*	*2.2*
SCANDANAVIAN										
Denmark	9	15	0.0	13.3	–	12	40	355	2,663	3.0
Finland	36	6	7.4	0.0	0.0	24	38	179	1,109	1.0
Norway	40	3	0.0	0.0	25.0	60	75	123	805	2.0
Sweden	40	5	0.0	13.5	100.0	12	40	304	3,040	1.0
*Group Mean**	*31*	*7*	*1.9*	*6.7*	*41.7*	*27*	*48*	*240*	*1,904*	*1.8*
SOCIALIST										
China	12	1	2.9	44.6	25.0	24	–	23	238	2.0
Czech Republic	1	2	0.0	0.0	–	84	125	25	300	3.0
Hungary	5	14	7.1	0.0	–	84	50	105	297	1.0
Poland	20	6	2.4	0.0	80.0	36	75	110	740	1.0
*Group Mean**	*10*	*6*	*3.1*	*11.2*	*52.5*	*57*	*83*	*66*	*394*	*1.8*

*The group mean equals the simple average of the country-level medians.

countries (this variable is missing for about one-half of our loan sample). Average maturity ranges from 12 months to 60 months. Ownership also varies considerably across borrower countries. For example, the median percentage of loans held by domestic banks is much lower in developing countries, presumably because of the relatively low level of financial development in these countries. In contrast, in the German legal origin countries, domestic bank ownership of loans predominates (e.g., 100% in Japan and 44% in Germany).

Table I also reports the creditor rights index score for each borrower country. Consistent with LLSV, countries with French or socialist legal origins tend to be associated with weak protection of creditor rights. Note that in our regressions (every table except for Table IV), we do not include either country-level or firm-level fixed effects because (i) there is (almost) no time variation in the key legal and institutional variables, (ii) for all firms located in the same country, they share the same country-level attributes, and (iii) we have one loan per firm in most cases. Thus, inclusion of either firm-level or country-level fixed effects would render it impossible to identify the country-level effects on loan terms. However, in order to examine how the interaction of creditor rights and borrower characteristics (size and tangibility) affect loan ownership and contract terms, we do perform firm fixed-effects regressions (Table IV) while dropping all country-level variables including creditor rights. Because we have many loans for each country but no variation in our key explanatory variables *within* countries, we cluster the error across all loans made in a single country to account for possible dependence in the error in *all* regressions.

B. Results

In our regression models, we begin by testing in reduced form how country-level variables affect ownership and both the price and nonprice terms of loans. We then examine how the impact of creditor rights on loans varies with borrower characteristics. Next, we consider how ownership variables interact with the loan contract terms. Note that these regressions cannot be viewed as pure reduced forms because the ownership and contract terms may be determined jointly. Thus, we are careful not to draw causal links between the two. Nevertheless, we think these results help illustrate the extent to which ownership may affect the ability of banks to enforce loan contracts. In robustness tests, we include additional institutional variables (legal formalism and information sharing registries) as explanatory variables. We also split the sample by whether a borrower country is developed or not, with additional tests performed on different subsamples and specifications.

Tables II through V report the main results. Table II contains the reduced-form models of ownership, and Table III the reduced-form models of loans' price and nonprice terms. In Table IV we interact creditor rights indexes with firm size and tangibility of assets, and in Table V we include loan ownership as explanatory variables.

Table II

Regressions Relating Loan Ownership to Firm, Loan, and Country Characteristics

We include but do not report coefficients on year indicators, percent change in GDP, industry indicators (two-digit SIC), indicators on loan type, purpose, and currency, and an indicator for firms located in Africa. Sovereign and firm ratings are converted to an index from zero to six as follows: 0 = unrated; 1 = Aaa; 2 = Aa; etc. In computing standard errors, we cluster by borrower country. The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The table reports coefficients, with t-statistics in parentheses.

	Log Number of Lenders		Share Held by Government Banks		Share Held by Domestic Banks	
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)
Creditors' rights	−0.067	−0.081	0.164	0.887	−5.143	−7.502
	(1.87)*	(2.40)**	(0.20)	(1.22)	(1.89)*	(3.78)**
One if French legal origin	0.071	0.132	0.694	0.374	−9.517	−3.079
	(0.63)	(1.22)	(0.27)	(0.21)	(1.52)	(0.76)
One if German legal origin	0.031	0.136	1.160	4.603	28.271	25.233
	(0.30)	(1.20)	(0.38)	(1.36)	(4.24)**	(5.07)**
One if Scandinavian legal origin	−0.256	−0.247	7.481	9.768	−10.739	−13.027
	(2.32)**	(2.31)**	(1.79)*	(2.41)**	(1.34)	(1.56)
One if socialist legal origin	0.122	0.093	7.238	4.194	−10.853	−2.902
	(0.84)	(0.53)	(3.24)**	(2.75)**	(1.08)	(0.52)
Log of market value of equity	0.056	0.053	−0.141	0.051	−3.272	−3.896
	(2.59)**	(2.34)**	(0.65)	(0.25)	(4.10)**	(5.04)**
Net income/assets	0.018	0.101	−2.137	−3.790	−12.456	−4.019
	(0.06)	(0.37)	(0.65)	(1.13)	(1.46)	(0.49)
Total debt/assets	0.173	0.125	4.433	3.655	−4.112	−6.246
	(1.60)	(1.38)	(1.53)	(1.56)	(1.15)	(1.43)
PP&E / assets	−0.070	−0.009	1.629	0.484	−13.659	−5.994
	(0.53)	(0.07)	(0.61)	(0.20)	(2.87)**	(1.59)
Firm rating	−0.091	−0.059	−1.195	−1.048	−2.634	−1.208
	(2.08)**	(1.50)	(1.65)*	(1.67)*	(1.94)*	(1.13)
One if rating is missing	−0.653	−0.520	−5.267	−4.948	−4.667	1.132
	(3.65)**	(3.30)**	(1.85)*	(2.13)**	(0.74)	(0.26)
One if multinational	0.134	0.145	0.824	0.847	−8.407	−7.492
	(2.01)**	(2.34)**	(0.78)	(0.86)	(2.36)**	(1.80)*
Sovereign rating	–	−0.126	–	−0.073	–	−7.286
	–	(3.53)**	–	(0.10)	–	(3.81)**
One if sovereign rating is missing	–	−0.467	–	−1.637	–	−30.971
	–	(2.30)**	–	(0.48)	–	(3.17)**
Log of GDP/capita	–	−0.110	–	−2.483	–	−1.952
	–	(1.86)*	–	(2.59)**	–	(0.74)
Private credit/GDP	–	0.002	–	−1.430	–	9.144
	–	(0.03)	–	(0.58)	–	(1.94)*
Estimation technique			OLS			
No. of observations	4,322	4,322	4,322	4,322	4,322	4,322
R-squared	0.24	0.25	0.08	0.10	0.38	0.41

*Significant at 10%; **significant at 5%.

B.1. Ownership

Consistent with the findings of Ongena and Smith (2000), we find that the concentration of loan ownership is higher, that is, the number of lenders is lower, in countries with better protection of creditor rights (Table II, columns 1 and 2). This suggests that diffuse ownership provides banks a tool to mitigate risk when their legal rights are weak. In contrast, loan ownership becomes less diffuse as the risk of the country, as measured by the sovereign debt rating, worsens (column 2).[19] With respect to firm characteristics, loan ownership becomes more concentrated when borrowers are small (the coefficient on the log of market capitalization is positive in columns 1 and 2) and opaque (the coefficient on the missing firm rating indicator is negative), consistent with earlier research on lending relationships (e.g., Petersen and Rajan (1994) for small firms).

Our empirical model has limited ability to explain variation in the share of loans held by government banks, although we do find government ownership to be higher in the formerly socialist countries and in the Scandinavian countries relative to the omitted category of countries with English legal origin (columns 3 and 4). Conversely, columns 5 and 6 of Table II indicate that the share of loans held by domestic banks is lower in the socialist and Scandinavian countries (statistically insignificant).

Legal protection of creditors is strongly associated with foreign banks' ownership of loans (columns 5 and 6). The effects are large economically as well as statistically. For example, a one standard deviation increase in creditor rights (an increase of about one of the index value) increases the share of foreign ownership (decreases share of domestic ownership) by five to seven percentage points. In addition, the share of loans held by domestic banks increases with total private credit to GDP (financial development), although this finding is true almost by definition. As noted earlier, we include financial development for consistency with the other results, and to demonstrate the robustness of the effects of creditor protection. With regard to firm characteristics, foreign banks hold a larger fraction of loans to large firms, to firms with more tangible assets (PP&E divided by assets), and to multinational firms.

B.2. Contract Terms

Table III reports the results for loan contract terms. We find first that the firm-level variables enter as expected: Large or profitable firms (net income divided by assets) borrow with less collateral, at longer maturities and with lower interest rates. High leverage is associated with higher interest rates, which could reflect both greater credit risk (lowering supply) and/or a higher level of loan demand. Firms with high levels of fixed/tangible assets (PP&E divided by assets) also borrow at longer maturities, all else being equal, presumably because the maturity of their assets is relatively long (e.g., real estate

[19] The negative coefficient on the indicator for a missing sovereign rating suggests that these countries have about the same loan concentration as loans made to firms in countries with a sovereign debt rating of Baa (coded as four).

Table III

Regressions Relating Loan Contract Terms to Firm, Loan, and Country Characteristics

We include but do not report coefficients on year indicators, percent change in GDP, industry indicators (two-digit SIC), indicators on loan type, purpose, and currency, and an indicator for firms located in Africa. Sovereign and firm ratings are converted to an index from zero to six as follows: 0 = unrated; 1 = Aaa; 2 = Aa; etc. In computing standard errors, we cluster by borrower country. The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The table reports coefficients (marginal probabilities in columns 1 and 2) with t-statistics in parentheses.

	Secured Indicator		Log of Maturity (Months)		Log of Drawn All-in Spread	
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)
Creditors' rights	0.006	0.014	0.119	0.093	−0.145	−0.100
	(0.43)	(1.10)	(4.47)**	(4.34)**	(2.04)**	(3.41)**
One if French legal origin	0.174	0.157	0.050	0.076	−0.196	−0.364
	(3.27)**	(2.97)**	(0.64)	(1.06)	(1.76)*	(5.00)**
One if German legal origin	−0.136	−0.089	−0.242	−0.112	−0.317	−0.335
	(4.34)**	(2.23)**	(2.88)**	(1.07)	(2.91)**	(4.87)**
One if Scandinavian legal origin	0.248	0.298	−0.161	−0.273	−0.768	−0.559
	(0.98)	(1.15)	(1.56)	(3.08)**	(4.47)**	(4.32)**
One if socialist legal origin	0.270	0.211	−0.124	−0.151	−0.883	−1.086
	(2.48)*	(1.95)*	(0.65)	(0.69)	(5.43)**	(5.55)**
Log of market value of equity	−0.035	−0.034	0.006	−0.002	−0.173	−0.153
	(4.34)**	(4.07)**	(0.38)	(0.17)	(5.59)**	(5.09)**
Net income/assets	−0.227	−0.212	0.330	0.451	−0.642	−0.825
	(1.85)*	(1.70)*	(2.21)**	(2.54)**	(1.23)	(1.60)
Total debt/assets	−0.069	−0.078	0.008	0.020	0.177	0.257
	(0.92)	(1.02)	(0.11)	(0.31)	(1.18)	(1.79)*
PP&E/assets	0.059	0.052	0.296	0.346	−0.028	−0.225
	(1.15)	(0.99)	(3.20)**	(3.85)**	(0.14)	(1.21)
Firm rating	0.025	0.028	0.062	0.088	0.194	0.155
	(0.67)	(0.75)	(1.30)	(1.80)*	(4.36)**	(3.31)**
One if rating is missing	0.064	0.065	0.323	0.455	0.779	0.639
	(0.52)	(0.54)	(1.63)	(2.29)**	(3.37)**	(2.88)**
One if multinational	0.050	0.051	−0.026	−0.021	−0.048	−0.008
	(2.20)**	(2.21)**	(0.56)	(0.54)	(0.46)	(0.10)
Sovereign rating	–	−0.016	–	−0.130	–	0.191
	–	(0.83)	–	(4.59)**	–	(2.94)**
One if sovereign rating missing	–	−0.064	–	−0.496	–	0.505
	–	(0.87)	–	(2.82)**	–	(1.82)
Log of GDP/capita	–	−0.025	–	−0.042	–	−0.214
	–	(1.36)	–	(1.07)	–	(1.35)
Private credit/GDP	–	−0.040	–	−0.130	–	0.427
	–	(1.48)	–	(1.99)	–	(5.49)**
Estimation technique	Probit		OLS		OLS	
No. of observations	1,643	1,632	3,792	3,792	1,255	1,255
R-squared	0.29	0.30	0.38	0.40	0.56	0.61

*Significant at 10%; **significant at 5%.

and equipment). However, we do not find that firms with more fixed assets are more likely to borrow on a secured basis, as one might expect if such assets represent "good" collateral. Safer borrowers (firms with a higher rating or a lower index value) obtain loans at lower interest rates, while opaque borrowers (firms without a rating) face higher costs. Indeed, the coefficients suggest that unrated borrowers pay interest rates similar to borrowers rated between Baa and Ba (coded as five).

Turning to legal protection, the focus of the paper, in columns 1 and 2 of Table III we find a positive association between creditor rights and collateral, although the coefficient is not statistically significant. Consistent with Davydenko and Franks (2005), collateral is used more frequently in the French legal origin countries and less frequently in the German legal origin countries relative to those with English legal origin. (Note that Davydenko and Franks compare loans made in France, Germany, and the U.K. only.)

Bank loan maturity appears to be especially sensitive to the legal environment (columns 3 and 4). Creditor rights are positively associated with maturity, and the coefficient is large both statistically and economically. A one standard deviation increase in creditor protection, for instance, comes with an increase in maturity of about 10%. Comparing extreme cases, a loan to a Mexican firm (where the index on creditors' rights equals zero) would have a maturity that is 40% shorter than a loan to a firm in the U.K. (where the index equals four). We also find that increased country risk (a lower sovereign rating, or a higher index value) is associated with shorter-term lending (column 4). These results suggest that maturity acts as a useful contracting tool when collateral is relatively ineffective (due to weak creditor rights protection) or when banks operate in environments in which the risk of government expropriation (a factor in a country's sovereign rating) is high. Maturity can be particularly useful in such environments because shortening the loan mitigates risk even absent well-functioning legal regimes and/or well-developed institutions (Diamond (2004)).

In the last two columns of Table III, we test how loan pricing reflects country-level legal and institutional variables.[20] Consistent with Davydenko and Franks (2005), we find that interest rates are higher in the English legal-origin countries than in those with French or German legal origin. We also find that greater financial development (as measured by a country's total private credit divided by GDP) is associated with higher interest rates (column 6). This seemingly surprising result is likely driven by loan *demand*, as a higher ratio of total private credit extended implies higher demand for loans and thus higher prices.

Most importantly, we find a negative and statistically significant relation between creditor rights and loan interest rates in both specifications (columns 5 and 6). The effect of creditor protection is economically large as well—a one

[20] Since most of the loans in the sample are float-rate loans and the interest rates are measured by the markup over LIBOR, term structure factors should not affect our results. As mentioned before, since interest rates on loans made to Asian countries are structured differently, we drop Asian loans in our interest rate regressions.

standard deviation increase in creditor rights is associated with a 10% to 15% decline in interest rates. Note that a country's sovereign debt rating has strong power to explain interest rates on loans, even more so than the firm's *own* credit rating. As expected, however, the impact of creditor rights falls by almost one-third when we control for the sovereign debt rating and the other outcome variables.

Taken together, the results from Tables II and III demonstrate that creditor rights have a significant impact on loan ownership as well as on the price and nonprice terms of loan contracts. In the presence of strong legal protection, creditors are more likely to concentrate their holdings, use collateral (weakly), lend on a longer-term basis, and charge lower rates. These results are consistent with theories of debt based on incomplete contracts and the transfer of control rights during bankruptcy (e.g., Aghion and Bolton (1992), Hart and Moore (1994, 1998)). How do loan contracts adjust to weaker legal protection? Loan ownership concentration falls (reducing the strategic default problem and facilitating diversification of risk across lenders), maturities shorten (effectively putting borrowers on a shorter leash), and interest rates rise (pricing the increased expected losses in default).

B.3. Interaction of Creditor Rights and Borrower Characteristics

Table IV reports firm fixed-effects regressions relating our ownership and contract variables to firm-level controls and the interactions of creditor rights with firm size (log of market capitalization) and asset tangibility (PP&E divided by assets). As discussed above, country-level variables are not identified in the models with firm fixed effects. The direct effects of firm-level characteristics, however, are weakly identified because we have some cases of multiple loans to the same borrower. Note that the effects of firm characteristics are estimated imprecisely because the fixed effects absorb most of the cross-firm variation.

We first find that the impact of creditor rights on loan ownership varies significantly with firm size. Together with the overall negative effect of creditor rights on the number of lenders (Table II), the positive interaction in Table IV (column 1, row 1) suggests that creditor rights affect ownership concentration most when borrowers are small. We find the opposite for ownership by foreign banks (column 3). Here, the interaction between firm size and creditor rights is negative, which is the *same* sign as the overall effect of creditor rights reported in Table II. Thus, creditor rights affect foreign ownership more for large borrowers than for small ones. This result is not surprising because foreign banks tend to lend mainly to large borrowers.

We also find that asset tangibility complements the effects of creditor rights on loan availability. Borrowers with high levels of tangible assets located in countries with strong creditor protection receive the lowest interest rates and the longest maturities on loans (columns 5 and 6). The reason for these better terms appears to be that collateral is more effective for such cases, as we find a positive and significant interaction between property, plant, and equipment and creditor rights. Thus, while the overall relation between creditor rights and

Table IV

Firm Fixed-Effects Regressions Relating Loan Contract and Ownership Terms to Firm-Country Interactions

We include but do not report coefficients on year indicators, percent change in GDP, industry indicators (two-digit SIC), indicators on loan type, purpose, and currency, and an indicator for firms located in Africa. Firm ratings are converted to an index from zero to six as follows: 0 = unrated; 1 = Aaa; 2 = Aa; etc. Country-level variables such as creditor rights are not identified in the model due to the inclusion of the firm fixed effects. For the model of secured lending, the dependent variable equals one if the loan is secured (estimated using a linear probability model). The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The table reports coefficients (marginal probabilities in column 4), with t-statistics in parentheses.

Explanatory Variables	Log Number of Lenders (1)	Share Held by Government Banks (2)	Share Held by Domestic Banks (3)	Secured Indicator (4)	Log of Maturity (5)	Log of All-in Spread (6)
Creditors' rights × log of market equity	0.020	−0.016	−0.603	−0.002	−0.001	−0.044
	(3.17)**	(0.17)	(3.44)**	(0.90)	(0.34)	(1.54)
Creditors' rights × PP&E/assets	−0.008	−0.952	4.468	0.163	0.308	−0.772
	(0.05)	(0.38)	(0.92)	(1.91)*	(2.42)**	(3.18)**
Log of market value of equity	0.077	−0.163	2.197	0.034	−0.035	−0.010
	(1.78)*	(0.26)	(1.82)*	(1.52)	(1.18)	(0.11)
Net income/assets	0.331	4.776	−12.539	0.449	−0.007	0.083
	(0.85)	(0.86)	(1.16)	(2.30)**	(0.03)	(0.22)
Total debt/assets	0.414	0.772	7.058	0.308	−0.168	1.036
	(1.67)*	(0.22)	(1.02)	(2.15)**	(0.96)	(4.13)**
PP&E/assets	0.002	2.931	6.288	−0.170	−0.644	1.586
	(0.00)	(0.37)	(0.40)	(0.58)	(1.53)	(2.18)**
Firm rating	−0.017	−0.683	0.391	0.008	−0.015	−0.050
	(0.23)	(0.65)	(0.19)	(0.14)	(0.31)	(0.94)
One if rating is missing	−0.239	−2.911	−2.109	0.093	0.013	−0.334
	(0.79)	(0.67)	(0.25)	(0.46)	(0.06)	(1.53)
Estimation technique	Borrower-Level Fixed Effects OLS					
No. of observations	4,322	4,322	4,322	1,749	3,792	1,255
R-squared (within)	0.09	0.04	0.08	0.12	0.22	0.37

*Significant at 10%; **significant at 5%.

collateral is positive but not statistically significant (Table III), the impact of creditor rights on the use of collateral is significantly stronger for borrowers that have more tangible assets, because creditor rights can make collateral more effective, and enhance loan availability, by working with the pledgeability (tangibility) of assets. These results are consistent with Braun (2003), who finds that industries with less tangible assets perform disproportionately worse when located in countries with weak legal protection and underdeveloped financial markets. We find that such firms that *do* borrow (i.e., those that appear in our data) borrow at high interest rates and at relatively short maturities.

B.4. Adding Ownership to Loan Contract Term Regressions

Table V introduces the ownership variables as right-hand-side regressors in the loan contract regressions. Adding these variables has little effect on the links between creditor rights and contract terms. Interestingly, loan ownership itself is significantly correlated with collateral, that is, loans with a greater number of participating banks (and less concentration) are more likely to be secured (columns 1 and 2). In particular, a one standard deviation increase in the number of participating banks is associated with a seven-percentage point increase in the probability that a loan is secured. This result suggests a *complementarity* between collateral (which raises the direct cost associated with default to borrowers) and diffuse ownership (which raises the deadweight costs associated with default, thereby preventing strategic default).

The participation of domestic banks is positively related to maturity (column 3 only) and negatively related to pricing (column 5 only), but these associations are absorbed by the addition of country-level outcome variables (columns 4 and 6). Combining results from Tables II, IV, and V, we conclude that foreign banks are more sensitive to the legal environment than domestic banks, and further that domestic banks have an advantage relative to foreign banks in overcoming information and control problems. Our results are also consistent with Mian (2006), who shows that foreign lenders are more likely to resolve defaulted loans in a formal court procedure than domestic lenders. He argues that foreign lenders are less able to establish a long-term relationship capital with borrowers, and hence have less relationship capital to exploit during private workouts; thus they resort to formal legal procedures. Our results are also consistent with Haselmann, Pistor, and Vig (2005), who find that foreign banks increase their lending volume to transition economies (in Central and Eastern Europe) substantially more than domestic banks following legal reform.

C. Robustness Tests

In all the robustness tests, we run model specifications that exclude the outcomes variables at the country level. We first introduce additional institutional variables to our models. Table VI reports the coefficients on the new variable indicated and on creditor rights; we suppress the other coefficients but these models include the same set of regressors as those in Tables II and III. In

Table V

Regressions Relating Loan Contract Terms to Country Characteristics (With Loan Ownership Variables Included as Regressors)

We include but do not report coefficients on year indicators, percent change in GDP, industry indicators (two-digit SIC), indicators on loan type, purpose, and currency, and an indicator for firms located in Africa. Sovereign and firm ratings are converted to an index from zero to six as follows: 0 = unrated; 1 = Aaa; 2 = Aa; etc. In computing standard errors, we cluster by borrower country. The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The table reports coefficients (marginal probabilities in columns 1 and 2), with t-statistics in parentheses.

	Secured Indicator		Log of Maturity		Log of All-in Spread	
Explanatory Variables	(1)	(2)	(3)	(4)	(5)	(6)
Log of number of lenders	0.039	0.039	0.018	0.006	−0.028	0.004
	(3.78)**	(4.18)**	(0.90)	(0.33)	(0.97)	(0.16)
Share held by government banks	0.001	0.001	0.001	0.002	0.004	0.002
	(1.23)	(1.12)	(0.69)	(0.95)	(1.12)	(0.86)
Share held by domestic banks	0.001	0.001	0.001	0.000	−0.004	−0.001
	(1.33)	(1.58)	(1.74)*	(0.72)	(4.11)**	(0.76)
Creditors' rights	0.012	0.022	0.126	0.096	−0.131	−0.095
	(0.94)	(1.88)*	(4.57)**	(4.15)**	(2.27)**	(3.34)**
One if French legal origin	0.169	0.138	0.057	0.076	−0.168	−0.341
	(3.17)**	(2.65)**	(0.72)	(1.06)	(1.88)*	(4.59)**
One if German legal origin	−0.145	−0.100	−0.272	−0.131	−0.264	−0.329
	(4.51)**	(2.69)**	(3.40)**	(1.23)	(2.96)**	(4.34)**
One if Scandinavian legal origin	0.309	0.364	−0.153	−0.277	−0.787	−0.567
	(1.14)	(1.33)	(1.40)	(2.99)**	(4.99)**	(4.66)**
One if socialist legal origin	0.288	0.212	−0.123	−0.158	−0.897	−1.077
	(2.47)**	(2.01)**	(0.61)	(0.71)	(6.22)**	(5.46)**
Log of market value of equity	−0.035	−0.034	0.008	−0.001	−0.179	−0.157
	(4.18)**	(3.87)**	(0.52)	(0.07)	(5.46)**	(4.83)**
Net income/assets	−0.221	−0.218	0.343	0.454	−0.566	−0.813
	(1.91)*	(1.85)*	(2.21)**	(2.52)**	(1.23)	(1.71)*
Total debt/assets	−0.062	−0.067	0.002	0.015	0.196	0.253
	(0.90)	(0.93)	(0.03)	(0.22)	(1.31)	(1.73)*
PP&E/assets	0.057	0.043	0.311	0.350	−0.103	−0.241
	(1.05)	(0.77)	(3.54)**	(4.01)**	(0.54)	(1.25)
Firm rating	0.022	0.024	0.068	0.090	0.191	0.156
	(0.64)	(0.68)	(1.43)	(1.85)*	(4.41)**	(3.45)**
One if rating missing	0.068	0.065	0.346	0.467	0.773	0.643
	(0.63)	(0.58)	(1.79)	(2.37)*	(3.44)**	(2.98)**
One if multinational	0.041	0.040	−0.021	−0.021	−0.062	−0.014
	(1.97)*	(1.93)*	(0.47)	(0.54)	(0.63)	(0.18)
Sovereign rating	–	−0.007	–	−0.126	–	0.184
	–	(0.39)	–	(4.45)**	–	(2.81)**
One if sovereign rating is missing	–	−0.045	–	−0.478	–	0.463
	–	(0.59)	–	(2.72)**	–	(1.66)
Log of GDP/capita	–	−0.020	–	−0.037	–	−0.207
	–	(1.07)	–	(0.94)	–	(1.30)
Private credit/GDP	–	−0.049	–	−0.132	–	0.404
	–	(1.83)*	–	(2.08)*	–	(4.85)**
Estimation technique	Probit		OLS		OLS	
No. of observations	1,643	1,632	3,792	3,792	1,255	1,255
R-squared	0.31	0.32	0.38	0.40	0.57	0.62

*Significant at 10%; **significant at 5%.

Table VI

Regressions Relating Loan Contract Terms to Firm, Loan, and Country Characteristics (With Other Institutions)

The regressions contain the same variables as those in columns (1), (3), and (5) of Tables II and III, although in this table we report only the coefficients on creditor rights and the additional institutional variable of interest. For the model of secured lending, the dependent variable equals one if the loan is secured, and it is estimated by Probit; marginal effects are reported. In computing standard errors, we cluster by borrower country. The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The table reports coefficients, with t-statistics in parentheses.

	Log Number of Lenders (1)	Share Held by Government Banks (2)	Share Held by Domestic Banks (3)	Secured Indicator (4)	Log of Maturity (5)	Log of All-in Spread (6)
Panel A: Include Legal Formalism						
Legal formalism	−0.015	2.735	−0.466	0.006	0.066	0.027
	(0.29)	(2.20)*	(0.14)	(0.29)	(1.28)	(0.26)
Creditor rights	−0.070	0.572	−5.206	0.008	0.129	−0.143
	(1.87)*	(0.74)	(1.85)*	(0.58)	(4.42)**	(1.89)*
No. of observations	4,321	4,321	4,321	1,643	3,791	1,255
R-squared	0.24	0.10	0.38	0.29	0.38	0.56
Panel B: Include Indicator for Countries with Private or Public Information-Sharing Registries						
One if country has information sharing	−0.015	0.179	−5.232	−0.065	−0.040	−0.027
	(0.16)	(0.09)	(1.03)	(1.38)	(0.36)	(0.19)
Creditor rights	−0.066	0.148	−4.673	0.010	0.122	−0.142
	(1.75)*	(0.17)	(1.73)*	(0.74)	(4.65)**	(2.10)*
No. of observations	4,322	4,322	4,322	1,643	3,792	1,255
R-squared	0.24	0.08	0.38	0.30	0.38	0.56

*Significant at 10%; **significant at 5%.

Panel A, we add legal formalism, which measures the costs of enforcing contracts through the courts (DLLS (2003)); in Panel B we add an indicator variable that equals one if a borrower country has either public or private information sharing registries. As discussed above, we do not view these institutional variables as outcomes variables such as economic and financial development; instead, we view them as possible determinants of the contracting environment between banks and firms. Information sharing mechanisms have been previously shown to be an important determinant of credit market development (e.g., Jappelli and Pagano (2002), DMS (2007)), and legal formalism has been shown to increase the time and monetary costs of adjudicating disputes. Overall, neither of these institutional variables has a significant impact on loan contracts, while their inclusion has only a very small effect on creditor rights.[21]

Next, we examine separately borrowers from developed versus developing countries (Tables VII and VIII). We use the World Bank definition of high-income economies to divide our sample. The developed country subsample includes Australia, Hong Kong, Ireland, Israel, New Zealand, Singapore, U.K., Belgium, France, Greece, Italy, Netherlands, Portugal, Spain, Austria, Germany, Japan, South Korea, Switzerland, Taiwan, Denmark, Finland, Norway, and Sweden. The rest of the countries are classified as developing countries.

The results of Tables VII and VIII suggest that most of our main results are not driven by either subsample. Standard errors tend to increase relative to the pooled results because we have only about 20 countries in each sample, and we cluster the residuals at the country level. We do find that the positive relationship between creditor rights and loan maturity seems strongest among the developed countries (Table VIII); in contrast, the coefficient is not statistically significant among the developing countries. Perhaps most striking, we find that the effect of creditor rights affects loan ownership differently across the two subsamples: Creditor protection affects government ownership positively for the developed countries but negatively for the developing countries, although the economic magnitude is fairly small in both sets of countries. In particular, a one standard deviation increase in creditor rights comes with an increase (decrease) of about three percentage points in government banks' share for developed (developing) countries. For foreign ownership, for which the coefficient magnitudes are larger, we find that foreign banks' participation in loans responds more to creditor rights in the developed countries.

Third, we report a series of additional robustness tests in which we vary both the sample and the specification. The result of greatest interest—the coefficients on creditor rights—is reported across each of these models in Table IX (Panels B–F). Panel A reproduces the coefficients from our baseline model (Tables II and III) to facilitate comparisons of the models' results. Note that in all of the models we include the same set of firm, loan, and country control variables as those included in the benchmark models, except where specifically indicated otherwise in the title to the panel.

Panels B through D report estimates for three industry subsamples, namely, manufacturing (SIC 2 or 3), transportation and utilities (SIC 4), and utilities (SIC 46, 48, or 49). We focus on the two industry groups (manufacturing and transportation and utilities) to examine whether the impact of creditor rights on loans is robust among firms with similar assets, profitability, and size. We also seek to further differentiate between transportation and utilities firms (we do not have enough observations to examine transportation firms alone). On the one hand, assets in transportation are potentially more pledgeable since they

[21] We also try splitting the information sharing indicators into two indicators, one for private and one for public registries, but neither of these is individually significant, nor are they jointly significant.

Table VII

Regressions Relating Loan Ownership to Firm, Loan, and Country Characteristics (Developed versus Developing Countries)

We include but do not report coefficients on year indicators, percent change in GDP, industry indicators (two-digit SIC), indicators on loan type, purpose, and currency, and an indicator for firms located in Africa. Sovereign and firm ratings are converted to an index from zero to six as follows: 0 = unrated; 1 = Aaa; 2 = Aa; etc. In computing standard errors, we cluster by borrower country. The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The developed county sample includes: Australia, Hong Kong, Ireland, Israel, New Zealand, Singapore, United Kingdom, Belgium, France, Greece, Italy, Netherlands, Portugal, Spain, Austria, Germany, Japan, South Korea, Switzerland, Taiwan, Denmark, Finland, Norway and Sweden. We define the other countries as developing (see Table I). The table reports coefficients, with t-statistics in parentheses.

	Log Number of Lenders		Share Held by Government Banks		Share Held by Domestic Banks	
Explanatory Variables	Developed (1)	Developing (2)	Developed (3)	Developing (4)	Developed (5)	Developing (6)
Creditors' rights	−0.067	0.067	2.871	−3.085	−10.736	−5.290
	(1.11)	(0.95)	(3.89)**	(2.66)**	(3.89)**	(1.82)*
One if French legal origin	0.032	0.285	4.168	−4.999	−7.145	−7.673
	(0.23)	(1.65)*	(2.54)**	(1.69)*	(1.57)	(1.49)
One if German legal origin	0.093	–	7.774	–	17.624	–
	(0.77)	–	(3.57)**	–	(3.81)**	–
One if Scandinavian legal origin	−0.225	–	14.054	–	−22.862	–
	(1.61)	–	(3.37)**	–	(2.21)*	–
One if socialist legal origin	–	0.455	–	1.706	–	−0.981
	–	(2.18)**	–	(0.48)	–	(0.15)
Log of market value of equity	0.046	0.084	−0.102	−0.509	−4.458	0.620
	(1.74)*	(5.69)**	(0.37)	(1.40)	(7.81)**	(1.05)
Net income/assets	−0.141	0.513	−2.189	0.763	−3.021	−31.602
	(0.51)	(0.83)	(0.57)	(0.11)	(0.40)	(2.16)*
Total debt/assets	0.156	0.195	−0.982	12.393	−5.020	−10.628
	(2.10)**	(0.76)	(0.64)	(3.57)**	(1.08)	(1.39)
PP&E/assets	0.077	−0.313	4.735	−1.454	−5.755	−21.241
	(0.66)	(1.36)	(1.75)*	(0.39)	(1.34)	(1.70)*
Firm rating	−0.077	−0.059	−0.704	−2.792	−1.895	−1.674
	(1.47)	(0.72)	(0.89)	(1.83)*	(1.42)	(0.78)
One if rating is missing	−0.663	−0.487	−3.682	−11.283	−3.116	−4.418
	(3.23)**	(1.36)	(1.31)	(1.55)	(0.52)	(0.45)
One if multinational	0.150	0.146	2.130	0.534	−11.043	−4.979
	(1.81)*	(1.16)	(1.89)*	(0.35)	(2.38)**	(1.31)
Estimation technique			OLS			
No. of observations	2,999	1,323	2,999	1,323	2,999	1,323
R-squared	0.27	0.29	0.14	0.15	0.43	0.42

*Significant at 10%; **significant at 5%.

Table VIII

Regressions Relating Loan Contract Terms to Firm, Loan, and Country Characteristics (Developed versus Developing Countries)

We include but do not report coefficients on year indicators, percent change in GDP, industry indicators (two-digit SIC), indicators on loan type, purpose, and currency, and an indicator for firms located in Africa. Sovereign and firm ratings are converted to an index from zero to six as follows: 0 = unrated; 1 = Aaa; 2 = Aa; etc. In computing standard errors, we cluster by borrower country. The data for loan characteristics come from LPC's Dealscan. The sample includes loans originated between 1994 and the middle of 2003. Firm characteristics come from Worldscope at the end of the year of loan origination. We drop loans to firms in financial industries, as well as public sector loans (SICs 6 and 9). The developed county sample includes: Australia, Hong Kong, Ireland, Israel, New Zealand, Singapore, United Kingdom, Belgium, France, Greece, Italy, Netherlands, Portugal, Spain, Austria, Germany, Japan, South Korea, Switzerland, Taiwan, Denmark, Finland, Norway and Sweden. We define the other countries as developing (see The table I). The table reports coefficients (marginal probabilities in columns 1 and 2), with t-statistics in parentheses.

	Secured Indicator		Log of Maturity		Log of Drawn All-in Spread	
Explanatory Variables	Developed (1)	Developing (2)	Developed (3)	Developing (4)	Developed (5)	Developing (6)
Creditors' rights	0.001	0.026	0.116	−0.050	0.002	−0.095
	(0.05)	(0.60)	(2.87)**	(0.92)	(0.03)	(0.95)
One if French legal origin	0.193	0.225	0.075	−0.159	−0.158	0.260
	(2.69)**	(2.75)**	(0.79)	(1.64)	(1.16)	(1.03)
One if German legal origin	−0.074	–	−0.258	–	−0.205	–
	(2.49)**	–	(1.75)*	–	(1.90)*	–
One if Scandinavian legal origin	0.190	–	−0.233	–	−0.469	–
	(0.85)	–	(1.85)*	–	(3.36)**	–
One if socialist legal origin	–	0.412	–	−0.166	–	−0.450
	–	(2.50)**	–	(0.80)	–	(1.86)*
Log of market value of equity	−0.029	−0.040	−0.004	0.007	−0.142	−0.147
	(4.65)**	(2.06)**	(0.20)	(0.43)	(5.49)**	(2.53)**
Net income/assets	−0.128	−0.317	0.175	0.908	−1.225	−0.011
	(1.07)	(1.30)	(1.09)	(1.95)*	(5.44)**	(0.02)
Total debt/assets	−0.028	−0.378	−0.046	0.156	0.304	0.317
	(0.44)	(6.17)**	(0.62)	(1.01)	(1.91)*	(0.70)
PP&E/assets	0.090	0.041	0.227	0.603	−0.086	−0.061
	(2.12)**	(0.44)	(2.76)**	(5.82)**	(0.54)	(0.30)
Firm rating	0.014	0.023	0.081	0.024	0.212	0.294
	(0.62)	(0.34)	(1.49)	(0.55)	(7.05)**	(2.98)**
One if rating is missing	0.001	0.146	0.331	0.234	0.841	1.601
	(0.01)	(0.60)	(1.48)	(1.11)	(4.75)**	(3.02)**
One if multinational	0.068	0.124	−0.006	−0.081	0.003	0.191
	(3.53)**	(2.73)**	(0.10)	(1.08)	(0.03)	(1.94)*
Estimation technique	Probit		OLS		OLS	
R-squared	1,002	567	2,639	1,153	854	401
No. of observations	0.39	0.30	0.44	0.32	0.64	0.66

*Significant at 10%; **significant at 5%.

are easier to be redeployed than those of utilities (e.g., aircrafts or trains versus power generators), in which case creditor rights might be more important for transportation firms than for utilities firms. On the other hand, utilities firms are more likely to be government protected or owned but they are perhaps the

Table IX

Regressions Relating Loan Terms to Country Characteristics (Robustness Tests)

This table reports the coefficients on the creditors' rights variable for various specifications and subsamples. Panel A reproduces the benchmark results from Tables II and III (Panel A). In the other panels, we vary either the sample or the specification. All regressions include the same set of explanatory variables as those reported in Tables II and III, columns (1), (3), and (5).

Explanatory Variables	Log Number of Lenders	Government Share	Domestic Share	Secured Indicator	Log of Maturity	Log of All-in Spread
Panel A: Full sample (Replication of Tables II and III)						
Creditors' rights	−0.067	0.164	−5.143	0.006	0.119	−0.145
	(1.87)*	(0.20)	(1.89)*	(0.43)	(4.47)**	(2.04)**
Panel B: Manufacturing Only (SIC = 2 or 3)						
Creditors' rights	−0.068	0.622	−5.009	−0.007	0.147	−0.172
	(1.76)*	(0.73)	(2.02)**	(0.34)	(6.77)**	(2.60)**
Panel C: Transportation and Utilities Only (SIC = 4)						
Creditors' rights	−0.132	−0.951	−7.164	0.020	0.142	−0.217
	(1.86)*	(0.77)	(2.04)**	(0.97)	(2.95)**	(2.59)**
Panel D: Utilities Only (SIC = 46, 48, or 49)						
Creditors' rights	−0.114	−1.273	−5.827	−0.002	0.215	−0.233
	(1.49)	(1.01)	(1.65)	(0.21)	(3.66)**	(2.49)**
Panel E: Drop Multinationals						
Creditors' rights	−0.022	−0.096	−6.190	0.011	0.124	−0.164
	(0.59)	(0.13)	(2.35)**	(0.72)	(4.10)**	(2.08)**
Panel F: Replace Current with 1-year Lags of Firm Characteristics						
Creditors' rights	−0.067	0.189	−4.783	0.017	0.112	−0.133
	(1.88)*	(0.23)	(1.80)*	(1.32)	(4.28)**	(1.89)*

*Significant at 10%; **significant at 5%.

only firms for which we know where exactly the assets are located and thus their governing laws.

The results are robust in each of these smaller samples. Specifically, we find a negative relationship between creditor rights and the number of lenders in each subsample (significant for two of three), a negative relationship between creditor rights and the share owned by domestic banks in each subsample (again, significant for two of three), and a significantly positive link from creditor rights to loan maturity and a significantly negative link from creditor rights to loan interest rates in all three samples. In fact, the impact of creditor rights on loan maturity and interest rate seems to be much greater for utilities firms than for the whole sample.

Approximately one-quarter of the firms in our sample have operations outside the borrower's home country. In such cases there may be a concern that collateral located outside the borrower's country may be difficult for lenders to repossess under the laws of the borrower's country. As noted earlier, we control for these multinational companies with an indicator. However, we also estimate our models without these observations. As shown in Table IX, Panel E, our results are largely unchanged. All of the signs on the coefficient remain the same, and in most cases the magnitudes increase somewhat. In the last panel of Table IX (Panel F), we replace concurrent values of the firm characteristics with lagged values because some of these ratios, particularly the leverage ratio, change when firms borrow. Again, we find very similar results to those reported in our main regressions.

As a final robustness test (not reported), we estimate an instrumental variables model, where the legal origin indicators serve as an instrument for creditor protection (similar to DLLS (2006)). The impact of creditor rights on the ownership and loan maturity remains similar to those reported in Tables II and III (both statistically and economically). The effect of creditor rights on secured lending remains statistically insignificant, and the effect of creditor rights on the interest rate loses statistical significance.

III. Conclusions

In this paper we test how creditor rights affect the terms of bank loans across a cross section of countries. Strong protection of creditor rights is associated with greater concentration of loan ownership, increased participation by foreign banks, longer-term lending, and lower interest rates. These results paint a clearer picture of how creditor protection generates good financial outcomes. With better protection, lenders can control borrowers' risk because they know they will be able to take assets, or to credibly threaten to take assets, ex post, in the event of default. Firms benefit from the resulting expansion of credit supply by borrowing at longer maturities and at lower interest rates. Absent strong legal protection, lenders resort to diffuse ownership and shorter maturities—two contracting tools that do not rely on laws or institutions—to control borrower risk.

Two caveats are worth repeating. First, our results reflect how laws and institutions affect financial contracts for relatively large firms only, in particular, those appearing in both Dealscan and Worldscope. Second, we are only able to explore the terms of loans actually made. Earlier and concurrent research focuses on the extensive margin, with evidence that stronger investor protection increases the total supply of credit. We study the intensive margin, and find that the terms of loans actually made vary with the legal and institutional environment.

Appendix: Brief Descriptions of All the Variables and Their Sources

A.1. Legal Variables

Variables	Description	Sources
Legal origin	Identifies the legal origin of the company law or commercial code of each country.	LLSV (1998)
Restrictions for going into reorganization	Equals one if the reorganization procedure imposes restrictions such as creditors consent; equals zero otherwise.	Bankruptcy and reorganization laws; LLSV (1998)
No automatic stay on secured assets	Equals one if the reorganization procedure does not impose an automatic stay on the assets of the firm on filing the reorganization petition. Automatic stay prevents secured creditors from gaining possession of their security. Equals zero if such a restriction does exist in the law.	Bankruptcy and reorganization laws; LLSV (1998)
Secured creditors first	Equals one if secured creditors are ranked first in the distribution of the proceeds that result from the disposition of the assets of a bankrupt firm. Equals zero if nonsecured creditors, such as the government and workers, are given absolute priority.	Bankruptcy and reorganization laws; LLSV (1998)
Management does not stay	Equals one when an official appointed by the court, or by the creditors, is responsible for the operation of the business during reorganization. Equivalently, this variable equals one if the debtor does not retain the administration of the property pending the resolution of the reorganization process. Equals zero otherwise.	Bankruptcy and reorganization laws; LLSV (1998)
Creditor rights index	An index aggregating different creditor rights. The index is formed by adding one when: (1) The country imposes restrictions such as creditors' consent or minimum dividends to file for reorganization; (2) secured creditors are able to gain possession of their security once the reorganization petition has been approved (no automatic stay); (3) secured creditors are ranked first in the distribution of the proceeds that result from the disposition of the assets of a bankrupt firm; and (4) the debtor does not retain the administration of the property pending the resolution of the reorganization. The index ranges from zero to four.	Bankruptcy and reorganization laws; LLSV (1998)

A.2. Institutional Variables

Variables	Definition	Sources
Check-based legal formalism	The index measures substantive and procedural statutory intervention in judicial cases at lower-level civil trial courts, and equals the sum of the following categories (each takes on the value of one or zero): (1) professionals vs. laymen; (2) written vs. oral elements; (3) legal justification; (4) statutory regulation of evidence; (5) control of superior review; (6) engagement formalities; and (7) independent procedural actions.	Survey of Lex Mundi/Lex Africa association of law firms; DLLS (2003); and details available at http://iicg.som.yale.edu/
(Moody's) Sovereign debt ratings	Rating agencies assess the risks of sovereign (government) bond default. These ratings therefore serve as a broad proxy for overall country risk. We create two indicators based on Moody's ratings on the long-term sovereign bonds for the borrower's country (denominated in dollars).	Moody's; information on the definitions of these ratings is available at https://www.bonddesk.com/moodys.html
Information sharing: Public registries	A public registry is defined as a database owned by public authorities (usually the central bank or banking supervisory authority), which collect information on the standing of borrowers in the financial system and make this information available to financial institutions. The variable is constructed in January of each year for every year from 1978 to 2003.	Survey of Banking Supervisors; Jappelli and Pagano (2002); World Bank (http://www.doingbusiness.org/ExploreTopics/GettingCredit/)
Information sharing: Private bureaus	A private bureau is a private commercial firm or nonprofit organization that maintains a database on the standing of borrowers in the financial system. Its primary role is to facilitate exchange of information among banks and financial institutions. The variable is constructed in January of each year for every year from 1978 to 2003.	Survey of Banking Supervisors; Jappelli and Pagano (2002); World Bank (http://www.doingbusiness.org/ExploreTopics/GettingCredit/)
Information sharing	Equals one if either a public registry or a private bureau operates in a borrower country, and zero otherwise. The variable is constructed in January of each year for every year from 1978 to 2003.	Survey of Banking Supervisors; Jappelli and Pagano (2002); World Bank (http://www.doingbusiness.org/ExploreTopics/GettingCredit/)

A.3. Miscellaneous Country-Level Variables

Variables	Definition	Sources
Economic development	Average of Log (per capita GDP) over the period 1994 to 2002; measures a country's overall economic status (developed vs. developing countries).	World Bank
Financial development	Claims on the private sector by banks as a share of GDP, averaged over the 1994 to 1997 period.	Demirgüç-Kunt and Levine (2001)
Percentage change in GDP	Percent change in GDP in two adjacent years, measuring the business cycle of a country.	World Bank

REFERENCES

Acemoglu, Daron, and Simon Johnson, 2005, Unbundling institutions, *Journal of Political Economy* 113, 949–995.

Aghion, Philippe, and Patrick Bolton, 1992, An incomplete contracts approach to financial contracting, *Review of Economic Studies* 59, 473–494.

Allen, Franklin, Jun Qian, and Meijun Qian, 2005, Law, finance and economic growth in China, *Journal of Financial Economics* 77, 57–116.

Bae, Kee-Hong, and Vidhan K. Ghoyal, 2004, Property rights protection and bank loan pricing, Working paper, Hong Kong University of Science and Technology.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine, 2003, Law and finance: Why does legal origin matter? *Journal of Comparative Economics* 31, 653–675.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine, 2005, Law and firms' access to finance, *American Law and Economics Review* 7, 211–252.

Benmelech, Efraim, 2006, Asset salability and debt maturity: Evidence from 19th century American railroads, *Review of Financial Studies* (forthcoming).

Benmelech, Efraim, Mark Garmaise, and Tobias Moskowitz, 2005, Do liquidation values affect financial contracts? Evidence from commercial loan contracts and zoning regulation, *Quarterly Journal of Economics* 120, 1121–1154.

Bergman, Nittai, and Daniel Nicolaievsky, 2007, Investor protection and the Coasian view, *Journal of Financial Economics* 84, 738–771.

Berkowitz, Jeremy, and Michelle White, 2004, Bankruptcy and small firms' access to credit, *RAND Journal of Economics* 35, 69–84.

Bolton, Patrick, and David Scharfstein, 1996, Optimal debt contracts and the number of creditors, *Journal of Political Economy* 104, 1–25.

Braun, Matias, 2003, Financial contractibility and asset hardness, Working paper, UCLA.

Claessens, Stijn, and Leora Klapper, 2005, Bankruptcies around the world: Explanations of its relative use, *American Law and Economics Review* 7, 253–283.

Davydenko, Sergei, and Julian Franks, 2005, Do bankruptcy codes matter? A study of defaults in France, Germany, and the U.K., *Journal of Finance* (forthcoming).

Demirgüç-Kunt, Asli, and Ross Levine, 2001, *Financial Structure and Economic Growth: Cross-country Comparisons of Banks, Markets, and Development* (MIT Press, Cambridge, MA).

Demirgüç-Kunt, Asli, and Vojislav Maksimovic, 1998, Law, finance and firm growth, *Journal of Finance* 53, 2107–2137.

Diamond, Douglas, 2004, Committing to commit: Short-term debt when enforcement is costly, AFA Presidential Address, *Journal of Finance* 59, 1447–1480.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2003, Courts, *Quarterly Journal of Economics* 118, 453–517.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2006, The law and economics of self-dealing, *Journal of Financial Economics* (forthcoming).

Djankov, Simeon, Caralee McLiesh, and Andrei Shleifer, 2007, Private credit in 129 countries, *Journal of Financial Economics* 84, 299–329.

Esty, Benjamin, 2004, When do foreign banks finance domestic projects? New evidence on the importance of legal and financial systems, Working paper, Harvard Business School.

Esty, Benjamin, and William L. Megginson, 2003, Creditor rights, enforcement, and debt ownership structure: Evidence from the global syndicated loan market, *Journal of Financial and Quantitative Analysis* 38, 37–59.

Giannetti, Mariassunta, 2003, Do better institutions mitigate agency problems? Evidence from corporate finance choices, *Journal of Financial and Quantitative Analysis* 38, 185–212.

Glaeser, Edward, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, 2004, Do institutions cause growth? *Journal of Economic Growth* 9, 271–303.

Hart, Oliver, and John Moore, 1994, A theory of debt based on the inalienability of human capital, *Quarterly Journal of Economics* 109, 841–879.

Hart, Oliver, and John Moore, 1998, Default and renegotiation: A dynamic model of debt, *Quarterly Journal of Economics* 113, 1–42.

Haselmann, Rainer, Katharina Pistor, and Vikrant Vig, 2005, How law affects lending, Working paper, Columbia University.

Jappelli, Tullio, and Marco Pagano, 1993, Information sharing in credit markets, *Journal of Finance* 48, 1693–1718.

Jappelli, Tullio, and Marco Pagano, 2002, Information sharing, lending and defaults: Cross-country evidence, *Journal of Banking and Finance* 26, 2017–2045.

Jayaratne, Jith, and Philip E. Strahan, 1996, The finance-growth nexus: Evidence from bank branch deregulation, *Quarterly Journal of Economics* 111, 639–670.

Kaplan, Steven, Frederic Martel, and Per Stromberg, 2006, How do legal differences and experience affect financial contracts? *Journal of Financial Intermediation* (forthcoming).

King, Robert, and Ross Levine, 1993, Finance and growth: Schumpeter might be right, *Quarterly Journal of Economics* 108, 717–738.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer, 1999, Corporate ownership around the world, *Journal of Finance* 54, 471–517.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 1997, Legal determinants of external finance, *Journal of Finance* 52, 1131–1150.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 1998, Law and finance, *Journal of Political Economy* 106, 1113–1155.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 1999, The quality of government, *Journal of Law, Economics, and Organization* 15, 222–279.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 2000, Investor protection and corporate governance, *Journal of Financial Economics* 58, 3–27.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny, 2002, Investor protection and corporate valuation, *Journal of Finance* 57, 1147–1170.

Lerner, Josh, and Antoinette Schoar, 2005, Does legal enforcement affect financial transactions? The contractual channel in private equity, *Quarterly Journal of Economics* 120, 223–246.

Levine, Ross, 1999, Law, finance, and economic growth, *Journal of Financial Intermediation* 8, 36–67.

McKinnon, R., 1973, *Money and Capital in Economic Development* (Brookings Institution Press, Washington, DC).

Mian, Atif, 2006, Distance constraints: The limits of foreign lending in poor economies, *Journal of Finance* 61, 1465–1505.

Norton, Joseph J., 1996, International syndicated lending: The legal context for economic development in Latin America, *Law and Business Review of the Americas* 3, 21–81.

Ongena, Steven, and David Smith, 2000, What determines the number of bank relationships? Cross-country evidence, *Journal of Financial Intermediation* 9, 26–56.

Petersen, Mitch, and Raghuram Rajan, 1994, The benefits of firm-creditor relationships: Evidence from small-business data, *Journal of Finance* 49, 3–37.

Posner, Richard A., 1973, *Economic Analysis of the Law* (Little-Brown, Boston, MA).

Rajan, Raghuram, and Luigi Zingales, 1998, Financial dependence and growth, *American Economic Review* 88, 559–587.

Rajan, Raghuram, and Luigi Zingales, 2003, The great reversals: The politics of financial development in the twentieth century, *Journal of Financial Economics* 69, 5–50.

Siegel, Jordan, 2005, Can foreign firms bond themselves effectively by renting U.S. securities laws? *Journal of Financial Economics* 79, 319–359.

Stiglitz, Joseph, and Andrew Weiss, 1981, Credit rationing in markets with imperfect information, *American Economic Review* 71, 393–410.

Strahan, Philip, 1999, Borrower risk and the price and nonprice terms of bank loans, Working paper, Banking Studies Function, Federal Reserve Bank of New York.

Stulz, Rene, and Rohan Williamson, 2003, Culture, openness and finance, *Journal of Financial Economics* 70, 261–300.

Sufi, Amir, 2007, Information asymmetry and financing arrangements: Evidence from syndicated loans, *Journal of Finance* 62, 629–668.

Westbrook, Jay Lawrence, 2000, A global solution to multinational default, *Michigan Law Review* 98, 2276–2328.