Industrial Competition as a Limit to Banking Market Power

Carlos Bellon*

Finance Department, Wharton School at the University of Pennsylvania†

November 13, 2011

Abstract

In this paper, I propose a new characterization of the relevant economic market for commercial lending. This definition takes into account an often-overlooked characteristic of the credit markets: the competitive nature of the borrowers' own business. As a result, I submit that the limit to how much rent can concentrated financiers extract from a particular borrower is determined by the strength of competition on *both* banking and industrial markets.

Using data from the Survey of Small Business Finances in a difference-in-differences test design, I find that in areas where banks are concentrated, firms that compete mostly within the banks' area of influence (and only those firms) face systematically higher interest rates than their peers. This effect is strong (60–70 basis points), but is restricted to a subset of the firm population. I interpret this empirical result as support for the claim that banks can only successfully exercise market (pricing) power over entire marketplaces, and not geographical areas or individual firms per se.

Keywords: Bank competition; Small business finance; Commercial loans.

*I am grateful to Itay Goldstein, David Musto, Nick Souleles and Bilge Yilmaz for their invaluable guidance. I would also like to thank Nickolay Gantchev and seminar participants at the University of Pennsylvania (Wharton).

†Address: 2300 Steinberg Hall-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104, USA, telephone: 267-237-1924, e-mail: *bellonc@wharton.upenn.edu*

1 Introduction

There is a long standing controversy surrounding the effects, if any, that bank competition has on the commercial loan market. Surveys of the empirical literature using the traditional structure-conduct-performance (SCP) paradigm of Mason (1939) and Bain (1951) find mixed evidence of the importance of bank concentration[1], while the question of whether bank competition can hamper the formation of beneficial relationships between lenders and borrowers that alleviate agency frictions (Petersen and Rajan (1995)) or encourage banks to shift away from transactional lending and into relationship loans (Boot and Thakor (2000)) is still unsettled.

My paper contributes to this literature by providing a new characterization of the relevant banking market based on the insights from vertical integration. This characterization allows me to define an innovative control group of firms unlikely to be affected by bank concentration. I then perform a difference-in-differences test to isolate the effects of bank competition on interest rates to small firms and present confirming evidence of the existence of overpricing by concentrated banks, if only to a subset of borrowers.

The main problem with empirical tests of bank's market power is one of identification: distinguishing the exercise of market power from the effect of unmeasurable characteristics of banks and/or firms that would lead to higher equilibrium interest rates or restricted access to finance is often difficult.

[1]Weiss (1989) finds significant (at the 5 percent level) positive association between interest rates and bank concentration on only 21 of the 47 datasets it reviews. Gilbert (1984) finds that 32 out of 44 studies surveyed report some evidence of association between market structure and bank performance, with only 25 showing statistically significant associations.

A possible solution to the identification problem is to use exogenous changes in the level of competition, such as those produced by mergers, to measure the effects of bank concentration. This literature strand is exemplified by Berger et al. (1998) and Sapienza (2002), that use bank level data to measure how mergers affect banks' loan portfolios and therefore credit availability.

Changes in banking regulation, such as the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994, are instrumental variables that can be successfully used in identifying effects of bank concentration. Zarutskie (2006) uses this setup to find that an increase in competition reduces access to credit for young firms, confirming the model of Petersen and Rajan (1995). More in line with this paper, Rice and Strahan (2010) use the Interstate Banking and Branching Efficiency Act of 1997 as an instrument to test how credit competition affect small firm financial decisions, and find that loosening branching restrictions leads to interest rate savings of 80 to 100 basis points. Their results agree with earlier work by Jayaratne and Strahan (1996) studying economic growth due to more efficient banking systems after pro-competitive intrastate branching reforms. Black and Strahan (2002) also use changes in regulation to both inter- and intrastate branching restrictions to measure a positive effect of bank competition on new incorporations, while Kerr and Nanda (2009), using a similar methodology find both increased firm creation and churning (where new startups fail within the first years following entry).

Alternative ways to deal with the identification problem are reviewed by Shaffer (2004) through the use of tests arising from the "new empirical industrial organization" literature, such as the revenue test developed by Rosse and Panzar (1977) and

the markup test of Bresnahan (1982).

Finally, Cetorelli and Strahan (2006) use the measure of dependence on external finance developed by Rajan and Zingales (1998) to determine treatment and control groups for a difference-in-differences test of whether bank concentration affects the distribution of small firm's size. It is this last methodology that I will employ in my empirical test, taking the control group from a new characterization of banking market definition.

Typically the boundaries of banking markets have been thought of geographically, set at the level of the MSA or county and the degree of competition in banking was measured by the number of banks that are "in range" of a given firm. In this paper, I argue that concentration should be measured over industrial markets, and these markets are in turn defined by the (in)dependence of their demand curves. Therefore, the relevant measure of financial competition is the number of banks that are in range of a given firm *or its competitors*. The fact that the industrial markets are interconnected makes the financial markets interconnected too.

I propose a new characterization of bank competition. One that takes into account an often overlooked characteristic of the credit markets: the competitive nature of the borrowers' own business. By drawing attention to the insights of the vertical relationship paradigm of industrial organization, I posit that the maximum surplus that can be extracted from small firms is limited not only by competition within the bank's own market, but also by competition amongst borrowers.

As a result of this characterization, I predict that firms that compete in wide geographical markets should be charged loan rates close to the competitive rate on

average, regardless of how concentrated their local banking markets are. Whereas, firms whose entire markets fall under the geographical area of influence of a concentrated banking market should experience significant overpricing in their loans. Therefore, firms that compete exclusively in local markets should be more affected by the exercise of banking market power than those with wider marketplaces.

If these results are correct, and there are two limits to a bank's ability to exercise market power, then only when *both* conditions are met (concentrated banking markets *and* local firms) will there be an impact in the equilibrium interest rate. In this research desing, identification is achieved by measuring the differential effects of a change in bank concentration (from concentrated to competitive banking markets) between firms that compete in wide geographical markets (the "control group") and those (local firms) whose competitors all lie within the area of influence of the banking market under study (the "treatment group"). This difference-in-differences setup controls for systematic variation between locations where banking is concentrated and those where it is competitive: if the only reasons for higher interest rates (after adjusting for known drivers of individual interest rates) are variation in firms' unobservable risk profiles or local credit market size, these effects would be captured by the main effects of either local status or bank concentration and the interaction effect (the "difference-in-differences" estimator) would be insignificant. On the other hand, if concentrated banks are indeed exercising market power and charging firms interest rates over and above the competitive rate, then such effects should be larger for firms that operate on a market whose geographical definition overlaps with the banking market (local firms), and this effect would show itself through a significant

coefficient in the interaction term.

I use data from the Survey of Small Business Finances (SSBF) that offers the advantage of letting me classify all firms in the survey according to relatively good proxies for the level of bank concentration *and* of industrial competition for a given geographical footprint, in this case whether or not they compete primarily in their local markets. With this information I can subdivide the survey population into four subcategories by cross tabulating the bank concentration and local competition variables and test for systematic differences in loans rates between these four subpopulations that cannot be attributed to the traditional determinants of loan pricing. The survey *directly* measures the geographical footprint of a company's business in the same units (MSA or county) used to calculate the degree of bank concentration, thus putting the cross-classification of observations on the same footing.

Confirming the tenants of my conceptual framework, I find that, everything else being equal, only firms whose competitive footprint overlaps with a concentrated banking area are significantly overcharged for financing. This surcharge is on the order of 60–70 basis points, and thus accounting for a third of the average spread (211 basis points, with a standard deviation of 14 basis points), or ten percent of the average interest rate (6.56 percent, with a standard deviation of 14 basis points). These results are robust to alternative specifications.

The key identifying condition in my methodology is that the only interaction between concentrated banking markets and local industrial markets relevant for loan pricing is the exercise of market power (no omitted interactions). I assume that, after accounting for the usual determinants of interest rates (control variables), systematic

differences between "treatment" and "control" groups (if any) are similar in competitive and concentrated banking markets. However, if this assumption does not hold two main kind of alternative hypotheses could explain the results of this paper: there could be an unobserved risk or cost factor systematically higher only for local firms in concentrated banking markets; alternatively, if indeed market power is being exercised, it could be the case that some kind of friction (different from industrial market structure) explains the higher interest rates that local firms experience. In Section 4, I discuss these alternative explanations at length and use the depth of information available in the SSBF to perform additional tests that support the no omitted interaction assumption.

An implication of the new characterization of banking markets raises an economic policy concern that was missing in previous discussions of bank concentration. Having no downstream competitor outside of the bank's reach, there is no industrial limit to the bank's market power, which is equal to the full monopoly rent in the industrial market. Thus, there is a strong incentive for oligopolistic financial institutions to enforce collusive pricing in the industrial (downstream) sector by setting high interest rates across the industry and expropriate most of the surplus generated, thus hurting consumers as they will perpetuate monopolistic pricing in the local downstream industries. Although this paper finds evidence for overcharging of local firms by concentrated banks, only conceptual conjectures can be made as to the effects on local industrial markets. There is a need for further research that could directly test whether in fact small firms that compete locally in areas where banks are concentrated charge prices closer to the collusive equilibrium than to the

competitive level.

The remainder of the paper is organized as follows. In Section 2, I lay the conceptual framework necessary to analyze the effects that competition amongst borrowers has on bank competition. In Section 3, I conduct empirical tests aimed at measuring the relative importance of geographical versus industrial competitiveness in determining loan rates for small firms. In Section 4, I analyze the robustness of the empirical tests of the previous section, discuss the plausibility of the model's key identifying assumption and its competing hypotheses. Section 5 concludes.

2 Conceptual framework

The industrial organization literature (see, for example Stigler and Sherwin (1985) and references therein) defines the market for a good as the area within which the price of that good tends to uniformity (taking into consideration transportation costs). This is the fundamental definition of Cournot and Marshall and it has been adopted by the Department of Justice and Federal Trade Commision[2]. The relevant market for antitrust analysis is based on "demand substitution factors, i.e., customers' ability and willingness to substitute away from one product to another in response to a price increase". This market may consist of a bounded geographical area (if it limits the customers' willingness or ability to substitute to other products) and/or a group of substitute products.

In the case of banking, the market definition currently in use is based on a 1963

[2]U.S. Department of Justice and Federal Trade Commission. Horizontal Merger Guidelines. August 19, 2010. Available online at http://www.justice.gov/atr/public/guidelines/hmg–2010.html

Supreme Court ruling (*United States v. Philadelphia National Bank*) which determines banking markets based on two elements. First, banking markets are geographically local, encompassing rural areas the size of a county and urban areas the size of a metropolitan statistical area (MSA) - a city and the suburbs around it. Second, the relevant product market is considered all services provided by commercial banks, including deposits, loans, etc. Despite this second dimension and the fact that most theoretical models of bank competition such as the Monti-Klein model[3] are models of *intermediation* in which characteristics of bank funding and lending play a role, the empirical literature on the consequences of bank concentration (see Amel and Starr-McCluer (2002) for a detailed review) has independently analyzed particular segments of banking services (deposits, consumer finance, commercial lending). I will continue that approach throughout the rest of this paper restricting my attention to the provision of commercial loans, drawing on the results of Adams et al. (2002) that provide theoretical support for the argument that empirical studies of market power that concentrate on either the input side or the output side, are not subject to significant misspecification error.

The geographical characterization of banking markets as local has seen strong empirical support based on surveys of customer behavior (see Gilbert (1984) for a general overview). For commercial loans, figures from the Survey on Small Business Finances (SSBF) show that, in 2003, 75% of all loans to small firms where arranged by banks less than 12 miles away from the firm's main office, and 90% by banks less than 32 miles away. There have been (sometimes strong) challenges in the

[3]Monti (1972), Klein (1971)

literature to this geographical definition of banking markets. Based on the above characterization of economic markets as the reciprocal side of price determination, many studies (see a review in Shaffer (2004)) have tried to analyze price behavior in different segments of the banking market to test whether they are well defined for antitrust purposes. The results are mixed, with some studies such as Berger and Hannan (1989) finding pricing evidence that the deposit market is local; and Hannan (1991) and Petersen and Rajan (1995) showing the same result for the commercial loan market. On the other hand, Jackson (1992) finds that at least some deposit markets should be defined nationally.

Yet, these challenges all present markets defined over *geographical* areas wider than the MSA as the alternative hypothesis. The stronger criticism to this market definition comes from the theory of contestable markets developed by Baumol et al. (1982), where the threat of entry into a concentrated market can induce competitive pricing even in the face of a monopolist incumbent. Under these conditions the monopolist bank will optimally charge its clients the highest rate that cannot be matched by competitor financiers; be it currently established competitors or potential new entrants (limit pricing as introduced by Bain (1949)), thus preserving its monopoly power. But in order for there to be any difference in pricing, the incumbent must enjoy some kind of privileged position vis-a-vis its financial competitors. It should then be the mechanism that allows a bank to keep its potential competitors at bay that determines how banking markets should be defined and which ones should be considered concentrated. There are two main potential sources of banking monopoly power identified by the literature: adverse selection and transportation

costs.

Dell'Ariccia et al. (1999) and Marquez (2002) argue that adverse selection represents a barrier to entry in banking: Information gathered through lending allows incumbent banks to better evaluate returning borrowers. Then, adverse selection stemming from a bank's inability to distinguish between borrowers seeking finance for untested projects and those rejected by the competition, generates a fixed cost that limits the number of competitors in the banking market. Dell'Ariccia (2001) expands this result, from a single period Bertrand setting, to a multi-period model of spatial competition. Hauswald and Marquez (2006) analyzes the mixed case in which the information generating process about a borrower is a decreasing function of the distance between bank and firm. Degryse and Ongena (2005), on the other hand, provide evidence supporting the theory that the source of monopoly power is a difference in costs[4] related to distance between lender and borrower and not asymmetry of information.

As a consequence of this analysis, the previous literature measures the degree of competition in banking by the number of banks that are "in range" of a given firm. Where a firm can be "in range" of a certain bank either because they are physically near each other or because the firm is not very informationally opaque to the bank.

My contribution to this literature is to draw attention to the fact that there is another factor influencing borrowers' ability to substitute away from the services of local financiers, and therefore another limit to the interest rate that even a monopolist bank can charge. The most a firm is willing to part with in the negotiation of funding

[4]Both relationship generation costs and monitoring costs could increase with distance, as they involve face to face meetings and similar hands-on interaction.

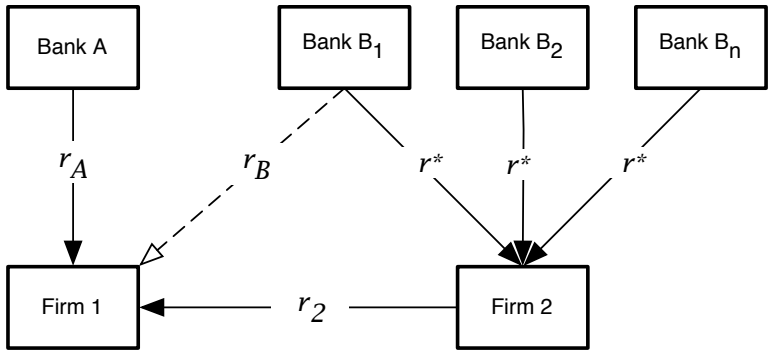
is its total surplus, less the risk-adjusted value of any outside option the entrepreneur might have. In short, a bank can extract rents from a firm only if the firm can extract rents from its customers. Previous literature takes this surplus as some exogenously determined quantity. But this is not a neutral simplifying assumption.

Building on the insights of the vertical relationship theory (see Tirole (1988)) we see that as industrial surplus is the result of competition in the firm's own industrial market it *can be influenced by the interest rates charged to its direct competitors*. Thus treating loans as projects independent of each other ignores a second channel through which banks compete. Competition amongst borrowers acts as a limit to the amount of overpricing that financial monopolists can successfully exercise. As long as a financial monopolist cannot influence its borrower's competitors this limit is set so that financing rates are not high enough to price the borrower out of its own industrial market.

The previous statement is true regardless of what the reason is for a particular financial intermediary to have a position of power when negotiating loan terms with a particular borrower: geographical concentration of banking, inside information monopoly or even legal barriers. Thus, even a bank that enjoys an incontestable monopoly position with respect to a certain firm is affected by the level of bank competition elsewhere, as long as this firm competes with others that do not fall under the influence of the monopolist financier. The fact that the industrial markets are interconnected makes the financial markets interconnected too.

As a consequence of this often overlooked feature of banking markets we have to amend the previous definition of banking concentration from the number of banks

Figure 1: Limits to a bank's market power



that are "in range" of a given firm, to the number of banks that are "in range" of a given firm *or its competitors*.

The upper bound to equilibrium interest rates is thus defined by the *minimum* of the limit price that will keep "foreign" banks away *and* the interest rate that will make the borrower loose its competitive race against "foreign" firms (in effect, killing the goose that lays the golden eggs).

I will try to clarify the issues above through the use of the following example (see Figure 1 and the Appendix for a more formal treatment). Let us assume a world in which there are two cities (A and B). In city A there is just one monopolist bank, whereas in city B there are a number of banks engaged in perfect competition with one another. Banks in city B charge firms located in this city the perfectly competitive rate, which is also the minimum rate that the monopolist bank in city A can charge its local borrowers[5] and break even.

If firms in city A do not compete with those in B the limit to how much the monopolist bank can charge is determined by r_B, the minimum rate that banks lo-

[5]For simplicity I will assume that firms and banks across locations are essentially identical.

cated in city B can charge firms located in A and still break even. This rate is in turn determined by whatever barriers sustain monopoly power (distance, information opaqueness, legal impediments to cross-border lending) and their strength determines the economic importance of bank concentration. Therefore, in this case firms in city A would be charged an interest rate significantly higher than the perfectly competitive one[6].

However, if firms in city A compete with those in B, even in the case with a high r_B and therefore high potential for rent extraction by the monopolist, bank A faces another limit to how high the interest rate it charges can be: competition from firms located in city B. Because banking markets in city B are perfectly competitive, B firms are charged the minimum possible interest rate and are can themselves charge low industrial prices that reflect this. Unless there are barriers to competition in the industrial market (transportation costs, switching costs, etc.) or firms in city B are themselves industrial monopolists, firms in city A will be forced to match the prices offered by their competitors in B and will retain little or no surplus for the monopolist bank to expropriate.

Because the limit to bank overcharging is determined by the *minimum* of two limits to competition; in order for geographically concentrated banks to successfully overcharge their local borrowers *both* limits must be significant. I argue that bank concentration can only be maintained (and thus should only be measured) over the industrial markets of borrowers (defined both by product choice and geographical

[6]Note that this would be the case even if there were many firms in city A all ferociously competing against each other as the monopolist bank can charge all firms a similarly high interest rate that forces them all to sell at the collusive price to break even and therefore expropriate all industrial surplus in the market.

reach).

It is not immediately apparent which of these two restrictions should be binding. The answer will hinge on how variable is the level of competition in the financial sector compared to the level of competition in the industrial sectors.

It is easy to see how, in a very competitive industrial market with low transportation costs, just one competitive local financial market would be able to discipline all others: low financial costs translate to low industrial prices in the competitive location, these disseminate through all interrelated markets and force competing firms to reduce their surplus and therefore the maximum amount they are able to pay local monopolist financiers through interest rates. In this case, even if local monopolist banks were able to erect unsurmountable barriers to forestall competition by "foreign" banks, they would not be able to exercise any market power, as there would be no surplus left to be extracted. Thus, we should expect to see very little overcharging of firms that compete in markets that exceed their bankers' geographical reach. Monopolist banks intent on profiting from their position would find it difficult to tap this source of potential profits unless banks can influence not just individual firms, but all their competitors in their own product market. The way to achieve this outcome is to overcharge firms that only compete with others in the bank's own location. In fact, these firms are likely to suffer from high levels of both limits to financial monopoly and therefore expect significant overpricing.

Theoretical prediction: Firms that compete in wide geographical markets should be charged loan rates close to the competitive rate on average, regardless of how concentrated their local banking markets are. Whereas, firms whose entire mar-

kets fall under the geographical area of influence of a concentrated banking market should experience significant overpricing in their loans.

Even tough the framework described in this section deals with a monopolistic bank, it is still applicable to situations where there is significant bank concentration in one location and significant competition in others. Of course, in theory bargaining mechanisms can be designed by which just two banks can compete away all their market power; but there are also other designs that allow them to successfully collude and charge the monopolistic price. Moreover, regulations limiting the size of individual firms and industries in a bank's total portfolio and borrower's concerns about sensitive information leaking to their competitors if they share the same bank make it more plausible that a few banks can successfully partition a loan market between them and act as this framework predicts.

This hypothesis raises an economic policy concern that was missing in previous discussions of bank concentration. If the effect here identified is strong enough to be felt empirically, then bank concentration may not only generate distortions in the loan market, but the industrial market as well, potentially leading to collusive pricing and loss of welfare for consumers. Having no competitor outside of the bank's reach, there is no industrial limit to the bank's market power, which is equal to the full monopoly rent in the industrial market. Thus, there is a strong incentive for oligopolistic financial institutions to enforce collusive pricing in the industrial (downstream) sector by setting high interest rates across the industry. Furthermore, the fact that the local concentrated banks are the only lenders to the entire local industry makes it very difficult for potential "long range" competitors to inform

themselves properly about the risks of the local industry and therefore face large potential winners' curse[7]. As we can see, concentrated banks have both the incentive to act as collusive devices for the local industrial markets and the ability to erect credible barriers to entry of other financiers to support this outcome.

3 Empirical design

The key proposition that this paper tests is whether concentrated banks can successfully exercise market power and charge some firms higher rates than they would have been charged in competitive banking markets.

The main problem with testing this hypothesis is distinguishing the exercise of market power from the effect of unmeasurable characteristics of banks and/or firms that would lead to higher equilibrium interest rates. For example, locations with a small supply of available funds would naturally result in a small number of banks in equilibrium and at the same time in higher funding costs, leading to high interest rates *and* high bank concentration indices, even if banks were in perfect competition. Alternatively, if firms in certain geographical areas are systematically riskier than those in other areas this could also lead to both a smaller and less profitable market for corporate loans (and therefore a smaller number of banks in equilibrium) and higher interest rates "fairly charged" to those firms.

[7]If one is to look at the problem of entry into an oligopolistic banking market, it is easy to see (as, for example, in Van Tassel (2006)) that, as long as there is underlying variation in the quality of entrepreneurs and banks learn about this quality by lending to them, incumbent banks will always have an advantage over foreign banks in lending to incumbent industrialists. The newcomer suffers from a winners' curse in that it is only able to lure clients away from informed "relationship bankers" by offering them rates well below their risk-adjusted expected return, otherwise the incumbent bank can always match any offer made by the new entrant.

In this paper I attempt to solve this problem and achieve identification by measuring the differential effects of a change in bank concentration (from concentrated to competitive banking markets) between firms that compete in wide geographical markets (the "control group") and those (local firms) whose competitors all lie within the area of influence of the banking market under study (the "treatment group"). This difference-in-differences setup controls for systematic variation between locations where banking is concentrated and those where it is competitive. Returning to the previous examples: if the only reasons for higher interest rates (after adjusting for known drivers of individual interest rates) are variation in firms' unobservable risk profiles or local credit market size, these effects would be captured by the main effects of either local status or bank concentration and the interaction effect (the "difference-in-differences" estimator) would be insignificant. On the other hand, if concentrated banks are indeed exercising market power and charging firms interest rates over and above the competitive rate, then (as discussed in Section 2) such effects should be larger for firms that operate on a market whose geographical definition overlaps with the banking market (local firms), and this effect would show itself through a significant coefficient in the interaction term.

The key identifying condition in this setup is that the only interaction between concentrated banking markets and local industrial markets relevant for loan pricing is the exercise of market power (no omitted interactions). I assume that, after accounting for the usual determinants of interest rates (control variables), systematic differences between "treatment" and "control" groups (if any) are similar in competitive and concentrated banking markets. If this assumption does not hold,

two alternative hypotheses could explain the results of this paper: there could be an unobserved risk or cost factor that is systematically higher only for local firms in concentrated banking markets; alternatively, if indeed market power is being exercised, it could be the case that some kind of friction (and not industrial market structure) explains the higher interest rates that local firms experience. In Section 4 I discuss these alternative explanations at length and use the depth of information available in the SSBF to perform additional tests that support the no omitted interaction assumption.

3.1 Data

In order to test the degree to which concentrated banks are able to successfully exercise market power, I use data on small firm financing. By focusing on small businesses we are less likely to face the confounding effect of the portfolio choice problem between different asset classes, as these firms are predominantly financed through debt (see Petersen and Rajan (1994)). Furthermore, the existence of banking market power is more feasible for small firms, as they are likely to face much stronger asymmetries of information and therefore make it much easier for banks to maintain local monopolies of information that can sustain significant overpricing. Also, small firms are more likely to face only a small number of competitors and thus are more likely to compete over regions that completely overlap the areas of influence of concentrated banks.

Data in this study is obtained from the 2003 Survey of Small Business Finances (SSBF). This survey has been conducted for the Board of Governors of

the Federal Reserve System in order to provide information about a representative sample of small businesses in the United States. The survey was first conducted in 1987 and repeated in 1993, 1998 and 2003. Complete documentation on the SSBF, including codebooks and detailed questionnaires, can be found at http://www.federalreserve.gov/Pubs/Oss/Oss3/nssbftoc.htm.

The data set is well known in the literature. Petersen and Rajan (1994) were the first to use data from the 1987 survey in order to study how banking relationships expand credit availability for small firms. Since then, the SSBF has been used on numerous occasions to study interrelation between banking market structure and small business finances (see, for example, Craig and Hardee (2007) or Cole et al. (1996)).

According to the survey's codebook, the target population for the survey are "all for-profit, non-financial, non-farm, non-subsidiary business enterprises that had fewer than 500 employees and were in operation as of year-end 2003 and on the date of the interview" (June - December 2004), representing a total population of 6.3 million small businesses. The 2003 SSBF that I use in this paper consists of a sample of 4,240 small businesses. Of these firms: 1,897 (44.7%) applied for a loan at some point during the three previous years; 1,757 (41.4%) were (eventually) approved; and 1,607 (37.9%) obtained a loan from an arms length institution (as opposed to a captive financial institution part of the borrower's group). This final group will constitute the population of interest for most of the empirical analysis in this paper. These most recently approved loans (MRAs) include lines of credit (1,053 observations, or 66% of the population of interest), capital leases (17 observations), mortgages (153 or

10% of the population) and other loans (381 observations, or 24% of the population).

The survey was designed as a stratified random sample, with over-sampling of the (relatively) larger firms (those with 20 to 499 employees). Sampling independently across strata ensures that strata are statistically independent; whereas the fact that observations that pertain to the same stratum are more homogeneous than the population as a whole can be exploited to produce more precise estimates. The 72 strata in the sample were generated by the cross-classification of three variables: number of employees, urban/rural status and census division, to ensure adequate representation of all subgroups.

As is the case for all surveys, there is some amount of missing data for most questions in the SSBF. The total incidence of missing data is about 1.8% of all values collected: with thirty percent of observations showing no missing values and 79 percent of observations having less than 3 percent of values missing (see the survey's codebook, p. 14). Although the public database contains multiple imputations for most of the missing values performed using a randomized regression model, it makes it possible to identify which answers have been imputed, and which are original answers. In this paper I only use actual answers in the estimation of all statistical models and drop observations with missing or imputed values.

The biggest advantage of the SSBF is the level of detailed information it provides on potential drivers of interest rates such as relationship, loan and firm characteristics, including matching information of bank characteristics and uses and sources of capital. This information allows me to determine characteristics not only of firms that got credit, but also of those that were rejected credit, as well as those that did

not want credit and make it possible to conduct ancillary tests to support the key identifying assumption. Furthermore, the survey also provides information about the geographical area over which a firm operates.

3.2 Variables

For the purposes of this paper, the SSBF offers the advantage of letting me classify all firms in the survey according to relatively good proxies for the level of bank concentration *and* of industrial competition for a given geographical footprint, in this case whether or not they compete primarily in their local markets. With this information I can subdivide the survey population into four subcategories by cross tabulating the bank concentration and local competition variables and test for systematic differences in loans rates between these four subpopulations that cannot be attributed to the traditional determinants of loan pricing.

In order to estimate the theoretical results from Section 2, I would like to mimic its setup as closely as possible. When it comes to the definition of the geographical unit that will play the part of location in my theoretical framework, the common unit of analysis traditionally used by the bank competition literature is the MSA or county (see, for example Hannan (1991)).

The commonly available measure of bank concentration is the Herfindahl index of commercial bank deposit concentration for the MSA or county where a firm's headquarters are located, derived from the FDIC Summary of Deposits data. The Herfindahl index is calculated as the sum of the squared market shares of all market participants multiplied by 10,000. The SSBF only reports this data categorically.

The survey reports whether the Herfindahl index is less than 1,000, between 1,000 and 1,800, or greater than 1,800.[8] It is this last category that I will consider corresponding to "concentrated" locations, whereas I will consider "competitive" locations those falling under the first two categories.

Once equipped with a measure of geographical bank concentration, ideally I would like to measure the strength and the geographical reach of industrial competition that firms face. In the terminology of the previous conceptual framework, I would like to know *where* the competitors of each firm are, how competitive the industrial marketplace is, how steep are transportation costs between locations and what is the *lowest* level of financial concentration that any competitor in the industry is exposed to, in order to calculate the industrial upper limit to loan prices. However this data is notoriously difficult to get hold of.

The SSBF provides us with a unique way of testing the effect on loan characteristics of industrial market structure. The survey *directly* measures the geographical footprint of a company's business. Firms were asked "Where does the business primarily sell or deliver its products or services?". This question was coded D3 in the questionnaire and allowed a total of nine possible answers in increasing order of geographical span, of which the first two ("within the city of the firm's main office" or "within the county/Metropolitan area of the firm's main office") are the ones I use to classify a firm as "local". The answer to this question refers to the same MSA (or county) over which the degree of bank concentration is calculated, thus putting the

[8]For illustration purposes, a market evenly split between five banks would score a Herfindahl Index of 2,000, whereas one evenly split between six banks would score a Herfindahl Index of 1,667 and one evenly split between ten banks would score 1,000.

cross-classification of observations on the same footing. More detail can be found in the survey questionnaire.

In the terminology of the Section 2, a "local" firm is one for which industrial transportation costs from other locations are either high or for which there are no out-of-town competitor firms at all (most likely because demand for different locations is independent of each other). Thus, it is a firm whose competitors all fall within the area of influence of the same geographically defined banking market. It is important to note here that the geographical footprint of a firm's sales is considered exogenous to conditions in the banking market. Even though a firm can freely choose where to sell its products, it cannot choose its competitors in the same manner. Industrial markets are defined along both a product and a geographical dimension. It is to this second dimension that I refer to when considering the geographical footprint of a firm as exogenous to conditions in the banking market, much as the type of products the firm sells is considered exogenous. It could, however be the case that a firm *chooses* to sell its goods only locally even if transportation costs are low and those goods should be considered a commodity and therefore have a wider geographical footprint as defined above. In this case, even if the firm only sells locally, its competitors do not, so it will face competition in its own local market from outsiders. In order to adjust for this case I constructed another variable that did not consider "local" any firm involved in the manufacture or sale of goods for which there is likely to be a wide geographical market (as measured by their SIC codes). However, there is no significant difference in the results of the analysis in this paper if this variable is used instead of the reported one.

Table 1: Description of subpopulations

	Concentrated banking market		Competitive banking market	
	Local firm	Non-local firm	Local firm	Non-local firm
Number of observations				
All respondents	996	1,075	1,099	1,064
Recently approved loans	313	468	386	485
Percentage of target population				
All respondents	25.8	22.2	29.4	22.7
Recently approved loans	23.1	23.5	29.1	24.3

3.3 Summary description of the data

Cross classifying bank concentration and local footprint, I end up with a total of four subpopulations. Table 1 tabulates the number of observations for each subpopulation, as well as the estimate of what percentage of the total survey target population each subpopulation represents. Note that the size of each of the four subpopulations is similar, both if we include all firms regardless of whether they asked for or obtained finance, and if we focus exclusively on those businesses with an approved loan in the previous three years.

The next step is to describe the key characteristics of the firms in each of the subpopulations and try to ascertain whether there are significant systematic differences between them that could account for the difference in spreads without recourse to the difference in market structures. In particular, the more similar the local and non-local firms are to each other, the more effective the latter will be as a control group. In panel A of Table 2 I tabulate the average levels of the main variables describing firm characteristics for all firms and for those with approved bank loans for each of

the four subpopulations: firms that compete in local markets, where there are very few local banks; firms that compete in local markets, where the local banking market is rather competitive; firms that conduct their business on a larger geographical scale than the county, where their local banking markets are concentrated; and firms that compete in wider-than-local markets, where local banking markets are competitive.

Most of the objective firm characteristics are quite similar between subpopulations, both if we look at the entire spectrum of firms or just focus our attention on those with recently approved loans. Although there appear to be significant differences in firm size and leverage between the subpopulations based on whether firms compete locally or not, these differences affect locations where banking markets are concentrated in the same way as those were banking markets are more competitive and persist whether or not firms get bank loans. Firms that compete in local markets are significantly smaller (under one half the size) and less levered than those that compete in regional or national markets.

Panels B and C of Table 2 underscore the fact that, apart from a significant difference in interest rates between local firms in concentrated banking markets, most loan characteristics are similar across subpopulations, as are the measures of the difficulty with which firms in all subpopulations can access finance. There are however significant differences that only apply to firms that compete in wide markets and are located in places with competitive banking markets. These firms are more often rejected credit, when they get it is shorter term and more likely to be secured and variable rate.

Table 2: Summary description of the data

	Concentrated banking market				Competitive banking market			
	Local firm		Not local		Local firm		Not local	
	Mean	Stdev.	Mean	Stdev.	Mean	Stdev.	Mean	Stdev.
Panel A: Firm characteristics								
Assets ($'000)								
All respondents	357	54	707	65	364	29	789	73
Recent loans	732	176	1,452	167	637	76	1,773	209
Average 3-year growth in sales (percent)								
All respondents	2.28	0.04	2.26	0.03	2.32	0.03	2.40	0.03
Recent loans	2.37	0.06	2.47	0.05	2.47	0.06	2.41	0.06
Firm age (years)								
All respondents	14.5	0.45	14.7	0.46	14.4	0.43	13.6	0.43
Recent loans	15.4	0.91	15.3	0.78	14.8	0.79	15.5	0.74
Firm leverage (total debt as a percentage of book value of assets)								
All respondents	95.7	41.6	176.7	86.3	79.8	11.3	178.0	50.2
Recent loans	51.9	4.2	96.5	35.8	103.3	15.6	90.0	24.5
Percentage of firms in an urban location								
All respondents	62.0	1.7	66.8	1.8	92.9	0.9	93.9	1.2
Recent loans	52.1	3.9	61.3	3.6	88.9	2.2	89.0	3.1
Panel B: Access to finance (percent)								
Rejected loan	5.4	0.9	4.6	0.8	4.0	0.7	6.6	1.1
New equity issued	4.1	0.8	5.8	1.0	5.5	0.9	7.2	1.1
Demanded any funds	44.7	2.0	47.9	2.1	44.7	1.9	52.0	2.1
Demanded debt	42.9	2.0	45.4	2.1	41.1	1.9	49.4	2.1
Panel C: Loan characteristics								
Maturity (months)	55.8	5.5	53.1	6.9	59.7	5.8	41.5	4.9
Percent long term	61.6	3.9	59.0	3.6	62.4	3.5	49.1	3.7
Amount ($'000)	195	54	335	44	232	40	557	90
Percent lines of credit	48.9	4.0	61.0	3.7	61.4	3.7	70.8	3.3
Percent secured	74.2	3.7	78.7	3.2	70.6	3.4	84.2	2.6
Fixed rate indicator	0.57	0.04	0.57	0.04	0.56	0.04	0.47	0.04
Interest (percent)	6.99	0.30	6.50	0.24	6.48	0.31	6.32	0.20
Spread (percent)	2.55	0.30	2.00	0.26	1.98	0.33	1.96	0.20
Fees (percent)	0.7	0.1	1.0	0.2	1.5	0.4	1.1	0.3

3.4 Regression specification

As noted above, in order to analyze the determinants of interest rates charged to small firms I use a "difference-in-differences" specification (see, for example Meyer (1995)). Taking advantage of economic theory to select a group of observations (the "control group") that should not be affected by changes in my variable of interest (the "treatment"). This allows me to control for potential unobservable differences between those observations that were and were not treated. The general specification is:

$$y_j = \alpha_0 + \alpha_1 Concentration_j + \alpha_2 Local_j + \beta Concentration_j * Local_j + \mathbf{Z}_j\delta + \epsilon_j$$

Where, j is an index across observed loans; $Local_j = 1$ if the firm competes only in local markets (thus belonging to the "treatment" group) and $Local_j = 0$ if it competes in wider geographical markets (and is therefore part of the "control" group). Finally, $Concentration_j$ is an indicator of bank concentration, equal to 0 if the observation is from an area where banking markets are competitive and equal to 1 if they are concentrated. The dependent variable, y_j, is the spread of the interest rate charged for the latest approved loan that firm j had (if any) in the past three years over the prevailing prime rate in the same month and year when the loan was approved[9], as reported by the Federal Reserve on its H15 Report. In the terminology of the Appendix this spread measures both the prevailing cost of capital for investments

[9]Other related papers (see, for example Rice and Strahan (2010)) use the interest rate as the dependent variable and the prime rate as one of the regressors. The results reported in this paper remain essentially the same if we were to adopt that specification, but I believe that the one chosen here makes interpretation of the main and interaction effects as drivers of the rate spread clearer.

with a risk profile comparable to that of a bank's better clients, and an estimate of the average operating costs of banks. Finally, all additional explanatory variables are captured by the vector $\mathbf{Z}_j$.

The purpose of this specification is to conduct a quasi-experiment to test what is the difference in interest rates that results from moving a firm with a geographically small footprint from a concentrated to a competitive banking market; and compare that difference to the one that arises from a similar move of a company that competes over a wider area. If the prediction from the Section 2 is correct, and there are two limits to a bank's ability to exercise market power, then only when *both* conditions are met (concentrated banking markets *and* local firms) will there be an impact in the equilibrium interest rate. This is a research design similar to Cetorelli and Strahan (2006), where they use the measure of dependence on external finance developed by Rajan and Zingales (1998) to determine treatment and control groups for a difference-in-differences test of whether bank concentration affects the distribution of small firm's size.

Several potential threats to the internal validity of the analysis (such as the effect of omitted variables, mismeasurement or trends in outcomes) are greatly reduced through this approach, as α_1 summarizes the way in which both groups ("treatment" and "control") are influenced by bank concentration and α_2 takes care of any systematic differences between both groups that are independent of whether firms are in a concentrated or a competitive banking market; leaving β as the true causal effect of the treatment on the outcome.

In general, the key identifying condition in this setup is that the only interac-

tion relevant for the independent variable is the treatment under study (no omitted interactions). Translated into my setting, the assumption is that the only interaction between $Concentrated = 1$ and $Local = 1$ relevant for loan pricing is the exercise of market power by concentrated banks. I assume that, after accounting for the usual determinants of interest rates (control variables), systematic differences between "treatment" and "control" groups (if any) are similar in competitive and concentrated banking markets. In other words, that in absence of treatment observations where $Local = 0$ and those where $Local = 1$ would have followed parallel paths.

However, the main identifying assumption of the empirical analysis cannot be itself tested and, if not true, there are alternative explanations for the test results. In Section 4, I present these alternative explanations, along with the results of ancillary tests that I am able to perform due to the richness of the survey dataset and lead me to believe they are less likely to account for the results of this paper.

In order to measure the effect that bank concentration elicits in equilibrium interest rates I need a good benchmark for what loan rates should have been under competitive markets in both the industrial and financial sectors. Therefore, following Petersen and Rajan (1995), Hale and Santos (2009) and Rice and Strahan (2010), I include a number of control variables in the specification of the empirical tests to account for all previously identified factors that might determine interest rates in equilibrium: macroeconomic environment variables, borrower controls, variables describing loan characteristics, relationship characteristics and variables describing the credit markets.

If financial markets are competitive, the most important factor determining variation in equilibrium interest rates would be the individual firm's risk profile. In order to control for the effect of differences in borrower characteristics that might impact riskiness or available industrial surplus, I include as regressors several measures of firm profitability: such as the firm's operating margin (EBIT/sales), return on assets (net income/assets) and turnover ratio (sales/assets). I also include the ratio of the firm's operating margin to the average operating margins of other firms in the same two-digit SIC code, as a way to measure the level of equilibrium surplus (and therefore industrial market power) that the firm enjoys. Furthermore, I adjust for the fixed effects of the firm's industry by including dummy variables for its SIC code. Regrettably for my purposes, the former measures may represent estimates of either a firm's riskiness, its efficiency or its market power depending upon the assumptions one makes about the precise nature of firm competition and are therefore not clear controls for idiosyncratic risk. My dataset only provides two variables that can reliably be considered to measure only riskiness: leverage ratio (total debt/assets) and a direct measure of the credit rating of the borrower derived from the Dun and Bradstreet credit score of the company: a number varying from one (safest) to five (riskiest).

Next, the issue of firm size is one of tremendous importance, as it may account for either pure bargaining power, riskiness, investment opportunities or information opacity: all of which are important factors underlying the dynamics of interest rates (for a more detailed exposition see the Appendix). In my test specification I measure firm size through the logarithm of the book value of assets and the number of

employees. I also consider the possibility that growing firms are subject to different pricing regimes that those under decline and add an indicator showing if sales have increased in the past year.

The last set of firm specific controls deals with the borrower's age. Following Petersen and Rajan (1995) I consider that concentrated banks may react differently from those in competitive markets to the uncertainty surrounding young firms. In order not to let this possibility interfere with the test, I control for the effect the logarithm of firm age has on the interest rate spread independently for firms in concentrated banking markets and in competitive ones.

The second driver of equilibrium interest rates in competitive markets are conditions in the capital markets that influence bank's operating and funding costs. The dependent variable in my regression, the interest rate *spread* over the prevailing premium rate at the time, adjusts for both average funding costs and an estimate of average bank operating costs. To isolate any remaining effects of market wide conditions, I control for the term structure of interest rates (calculated as the spread of the ten year Treasury Bond over the three month T-Bill) and the credit rating spread (calculated as the difference between the average interest rate on a BBB bond over a AAA bond).

There is a large literature on the effects that relationships between borrower and financier can have on loan characteristics (see Boot (2000) for an excellent review). A financier can acquire proprietary and reusable information through the course of repeatedly lending to the same firm helping to alleviate adverse selection and agency problems and reducing equilibrium interest rates. In an effort to measure

the strength of each player's bargaining position, I include in my model the length of the relationship (in years) between lender and borrower and whether the lender is considered by the borrower to be its "primary" financier. But intensity is not the only relevant measure of relationship for loan pricing. Degryse and Van Cayseele (2000) find that broader relationships in which the financier has access to other informative services with the borrower (cash management, credit card processing, etc.) reduce asymmetries of information and are reflected in lower rates. To measure the breadth of relationships I include three indicator variables in my model that measure whether the firm has deposits, information and/or non-information services with the institution that approved the loan.

But there is also a cost associated with relationship banking: the potential hold-up problem stemming from the relationship bank's information monopoly (see Rajan (1992)). Multiple relationships can alleviate this potential hold-up problem affecting interest rates. To control for this effect, I look at all relationships that the borrower has with other financiers: I measure the *total* number of relationships that the borrower has with financial institutions and the number of *lending* relationships. In order to put these into context, I also control for the longest lasting relationship the borrower currently has, and the distance to the furthest removed financier with which the borrower has a relationship (that indicates the potential area over which the borrower has some bargaining power). Also, the characteristics of the lender can determine different pricing systems. To account for this, I include indicator variables for whether the financier is a bank, a non-bank financial institution (such as an insurance or a leasing company), or other (individual, venture capitalist, etc.).

Finally, I look at the application process itself, measuring the number of times the borrower applied for a loan, the number of loan renewals, an indicator of whether the borrower was ever denied a loan in the past three years and another indicator of whether it did not apply for a loan for fear of refusal. All these variables can point to, otherwise unobservable, borrower quality concerns, as well as self-selection issues. Also, because distinct corporate governance structures may signal (or result from) agency problems, I control for their possible effect on equilibrium interest rates through indicators that measure if the firm is a corporation (including S-corporations), if it is owner-managed or if it is family owned. I also control for the effects of geographical location by including an indicator for whether the headquarters of banks and firms lie in rural or urban areas and nine regional dummies.

The conditions associated with the provision of capital span several dimensions (amount, maturity, security, etc.) and alternative funding offers might differ not just in the rate, but in any of the other dimensions. Therefore I include a number of regressors in the model that account for the more salient loan characteristics likely to result in different pricing regimes. Failing to control for these characteristics could interfere with my empirical test if they have a significant impact on pricing and are not evenly distributed amongst subpopulations. In our case, for example, the percentage of loans that are secured (by collateral, compensating balances, etc.) is significantly higher for non-local firms in competitive banking markets than other subpopulations (see Table 2, panel C). On the other hand, if we include as explanatory variables loan characteristics that are likely to be determined simultaneously with interest rates we could significantly bias our estimates. In the base specification of my model I have

therefore only included a few key loan characteristics in the base specification that may determine different pricing regimes but can be reasonable construed as exogenous. These include an indicator showing whether the agreed upon interest rate was fixed or variable and a variable indicating whether the loan was a renewal or a *de novo* loan.

Several theoretical papers (see, for example, Sengupta (2007) and references therein) have pointed out that by adjusting the contractual structure of loans firms can not only change their inherent risk profiles, but also signal their quality to potential financiers and thus influence their equilibrium interest rates. In order to control for this, I include dummy variables that control for the type of loan (line of credit, capital lease, mortgage, vehicle loan, equipment loan or other), for whether or not there was any kind of security attached to the loan (collateral, compensating balance or personal guarantee) and 6 dummy variables indicating the presence of each type of possible collateral (inventory or accounts receivable, business equipment or vehicles, business securities, business real estate, personal or other assets). The inclusion of these variables in the specification is more likely to lead to simultaneous equation bias. However, the results presented in Section 4 (where I estimate a model with no loan characteristics as regressors) show that no appreciable bias is incurred in.

3.5 Test results

The population of interest on which the regression is estimated is comprised of those companies that have some kind of loan (including lines of credit) approved in the three years ending December 2003 (1,757 or 41.4% of observations). Firms with

loans from captive financiers are dropped from the sample (86 or 2% of observations) as are those whose area of competition does not fit into the theoretical framework[10] (42 or 1% of observations).

I estimate the regression above using weighted least squares, taking the sampling weights provided by the SSBF that account for oversampling of larger firms and unit non-response. Since there may be a common element to the regression error across firms in each of the four subgroups I cluster standard errors by subgroup. The oversampling of (relatively) larger firms occurs to ensure there are sufficient numbers in the sample for analysis of this group; because they are a small percentage of the target population of small businesses, but are of special interest to researchers.

The first column of Table 3 reports the benchmark regression of this paper, linking the interest rate spread paid on the most recent loan to the level of bank concentration. The key metric we are concerned with is the coefficient on the interaction between the indicator variable for bank concentration and the one for local competition. As predicted by our main hypothesis this coefficient is statistically significant at the five percent level and economically important. This leads us to conclude that indeed small firms whose competitive footprint overlaps that of geographically concentrated banking markets are being "overcharged" relative to fully competitive "fairly priced" loans. The interaction effect is economically significant, entering the specification at 69 basis points, and thus accounting for a third of the average spread (211 basis points, with a standard deviation of 14 basis points), or ten percent of the average interest rate (6.56 percent, with a standard deviation of 14 basis points).

[10]Firms that primarily sold their goods outside USA (32), over the internet (4) and those that answered "other" (6).

Table 3: Model Specification and Estimation

This table reports regressions of the interest rate spread over the most recent loan on bank concentration and other borrower, loan and environmental characteristics, for both the treatment (local firms) and control groups (non-local firms). Several alternative specifications of the difference-in-differences test are shown using different sets of covariates provided by the previous literature. Model I is my base model. Model II is a more parsimonious specification without loan characteristics as regressors. Model III follows the specification of Rice and Strahan (2010). Model IV follows Petersen and Rajan (1994), while Model V follows Petersen and Rajan (1995).

Models	I	II	III	IV	V
Panel A: Main and interaction effects					
Concentrated	1.33	1.69	0.20	0.21	1.09
banking market	(2.09)	(2.00)	(2.05)	(2.31)	(2.20)
Does firm compete	-0.35^{*}	-0.16	-0.40^{**}	-0.33^{**}	-0.32^{*}
in local markets?	(-2.51)	(-1.52)	(-4.43)	(-5.08)	(-2.61)
Interaction effect	0.69^{**}	0.53^{**}	0.62^{***}	0.76^{***}	0.54^{***}
	(4.14)	(4.44)	(24.57)	(9.70)	(9.18)
Panel B: Environmental factors					
Term structure premium	0.96^{**}	0.94^{**}	1.05^{**}	1.01^{**}	1.03^{**}
	(4.42)	(4.19)	(4.04)	(3.97)	(4.31)
Default premium	1.06	1.34^{*}	0.50	0.79	0.68
	(1.89)	(2.92)	(1.22)	(1.66)	(1.17)
Indicator if borrower	0.37	0.39	0.50	0.53	0.46
is in urban loation	(1.15)	(1.47)	(1.49)	(1.22)	(1.22)
Panel C: Firm characteristics					
Indicator if borrower	-0.23	-0.24	-0.35	-0.38	-0.38
is a corporation	(-0.89)	(-0.77)	(-0.91)	(-1.03)	(-1.01)
Indicator if borrower	-0.32	-0.28			
is family owned	(-0.70)	(-0.73)			
Log of firm assets	-0.26^{*}	-0.38^{**}	-0.26^{**}	-0.29^{*}	-0.23^{*}
	(-2.54)	(-4.68)	(-3.88)	(-2.99)	(-2.53)
Number of employees	-0.34	-0.38			
('00)	(-1.61)	(-2.00)			

t-statistics in parenthesis. Significance denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

(Continues on next page)

Models	I	II	III	IV	V
(Continued from previous page)					
Indicator if sales	0.16	0.09		0.11	
are stable	(0.32)	(0.17)		(0.21)	
Indicator if sales	0.34**	0.18		0.19	
are growing	(3.80)	(0.95)		(1.93)	
Leverage	0.02	−0.01			0.00
	(0.55)	(−0.35)			(0.07)
Total debt				0.04	
($ mn)				(0.22)	
EBIT Margin	−0.45	−0.37		−0.22	
	(−1.62)	(−1.23)		(−0.77)	
Margin relative to	0.01	0.01			
industry peers	(1.25)	(1.27)			
ROA	−0.01	0.01	−0.01		
	(−0.45)	(0.56)	(−0.48)		
Turnover	−0.01	−0.02**			
	(−2.20)	(−3.31)			
Firm Age	−0.02	−0.01	−0.02**	−0.02*	
(years)	(−0.72)	(−0.75)	(−3.26)	(−2.78)	
Log of age if banking	0.22	0.30			−0.09
is competitive	(1.32)	(1.41)			(−1.99)
Log of age if banking	−0.19	−0.20			−0.44*
is concentrated	(−0.53)	(−0.49)			(−2.37)
Panel D: Relationship characteristics					
Number of institutions	0.03	0.04	0.13		
with relationship	(0.23)	(0.43)	(0.99)		
Number of lenders	0.08	0.03		0.13	0.13
	(0.81)	(0.31)		(1.06)	(0.89)
Borrower has deposits	−0.34	−0.30	−0.44	−0.45	−0.41
with lender	(−1.27)	(−0.81)	(−1.41)	(−1.35)	(−1.59)
Borrower has informative	−0.23	−0.59	−0.43	−0.45	−0.41
services with lender	(−1.07)	(−1.67)	(−1.31)	(−1.03)	(−1.36)
Borrower has no-info	0.22	0.41	0.80**	0.73**	0.69***
services with lender	(1.18)	(1.66)	(4.37)	(4.14)	(6.62)

t-statistics in parenthesis. Significance denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

(Continues on next page)

Models	I	II	III	IV	V
(Continued from previous page)					
Relationship length	0.01	0.02	0.01	0.01	0.01
(years)	(0.63)	(1.01)	(1.55)	(2.10)	(1.33)
Indicator if lender	0.43**	0.21			
is primary institution	(5.13)	(1.10)			
Longest relationship	−0.00	−0.00			
(years)	(−0.05)	(−0.22)			
Furthest relationship	0.16	0.25***			
('000 miles)	(0.92)	(6.07)			
Indicator if borrower	1.73	1.32			
has been denied credit	(1.47)	(1.03)			
Indicator if borrower	0.58	0.53			
feared denial	(1.00)	(1.17)			
Number of times	−0.03	−0.03			
applied for a loan	(−0.52)	(−1.35)			
Number of renewals	0.07	0.12*			
	(1.78)	(2.90)			
Panel E: Loan characteristics					
Non-bank Financial	2.83**	3.31**	3.24**	3.19*	3.30**
Lender	(3.24)	(3.85)	(3.25)	(3.01)	(3.62)
Non-financial	−0.80	−0.33	−0.81	−0.78	−0.72
Lender	(−0.86)	(−0.47)	(−0.59)	(−0.88)	(−0.58)
Indicator if interest	1.54***		1.41***	1.39***	1.42***
rate is fixed	(9.05)		(18.35)	(9.19)	(13.06)
Indicator if loan	0.58*				
is a renewal	(2.40)				
Indicator if loan	−1.21**				
is secured	(−5.18)				
Panel F: Fixed effect dummy variables					
SIC codes (9)	*yes*	*yes*	*yes*	*yes*	*yes*
Census division (9)	*yes*	*yes*	*yes*	*yes*	*yes*
D&B scores (5)	*yes*	*yes*	*yes*	*no*	*yes*
Collateral type (7)	*yes*	*no*	*no*	*no*	*yes*

t-statistics in parenthesis. Significance denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

(Continues on next page)

Models	I	II	III	IV	V
(Continued from previous page)					
Loan types (6)	*yes*	*no*	*no*	*no*	*no*
Observations	1311	1320	1375	1338	1402
R^2	0.561	0.504	0.509	0.507	0.515

t-statistics in parenthesis. Significance denoted by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

If we center our attention on the main effect of bank concentration, we see that its size is large (around 150 basis points) but not statistically significant; whereas the effect of competing only in local markets is negative, somehow significant (at the 10 percent level) and of a smaller magnitude (30 basis points). Neither of these results are very strong basis for any significant inference, but the sign of the main coefficient on local firms points to the fact that uncertainty about demand and competition (shown in the effect of selling products in local markets that are easier to monitor) may have a larger impact on pricing than the increase in risk concentration that comes from selling goods only in local markets.

As expected, there appear to be a number of different pricing regimes: dummy variables for whether the rate was fixed or variable, whether the loan was a renewal and whether there was any security involved are all significant and economically important (entering the specification at between 60 and 150 basis points). Other critically important variables are those related to the nature of the financier. Loans issued by non-banks are economically (283 basis points) and statistically significantly (at the five percent level) more expensive than those issued by banks.

What may be more puzzling from the results in Table 3 is the fact that, with the exception of asset size, the term structure premium and sales growth, very few

other variables appear to have a statistically significant impact on the spread[11]. In particular, in terms of relationship variables, only primary status seems to have a significant impact on prices. Loans obtained from the self-described primary bank are 43 basis points more expensive than those obtained from other institutions (again statistically significant at the five percent level). This is consistent with the fact that the previous literature has often found contradictory effects in when estimating the effects of relationships on interest rates. For example, Berger and Udell (1995) find small firms with long banking relationships pay lower interest rates on their lines of credit, whereas Petersen and Rajan (1994) cannot find a statistically significant effect of relationship length on interest rates.

4 Robustness and alternative explanations

4.1 Robustness

To demonstrate the robustness of my results I estimate several variations of the baseline model (Model I, in Table 3). Specifically, I fit a more parsimonious specification (Model II) without any loan variables as regressors to prevent any possible endogeneity biases derived from the fact that many loan characteristics may be determined simultaneously with rates. I also estimate a series of models based on the specifications of the most influential papers in the literature that analyze the effect of bank concentration on interest rates using SSBF data. Model III follows the specification of Rice and Strahan (2010) as closely as possible with the public

[11] Rice and Strahan (2010) also finds no consistently significant effect of most of their control variables and in particular, relationship variables.

data available. Model IV follows the specification of Petersen and Rajan (1994), while Model V follows Petersen and Rajan (1995). These models include fewer relationship variables and those dealing with firm characteristics, they are also missing important loan characteristics such as whether the loan is a renewal. The key difference between Model IV and V is that the latter allows for different effects for firm age to influence interest rates. Finally, Table 3 shows that the sign and magnitude of the coefficients of interest does not change substantially across model specifications.

The main threat to the validity of the difference-in-differences research design used in this paper is the possibility of an interaction between $Local = 1$ and $Concentration = 1$ other than the one under study (omitted interactions). If this were the case, there would have to be a factor or characteristic specific to firms operating in local markets in areas where banking is concentrated (and only to them) that would result in higher equilibrium interest rates.

Ideally, observations would be assigned randomly to the treatment or control groups therefore ensuring that no other systematic force was driving the results. Regrettably we cannot perform such an experiment. In such cases, the traditional approach is to study the lack of omitted interactions based on whether or not both groups are similar in the observable characteristics. I will argue below that this is the case in our setting and perform tests that show that those variables, which are indeed systematically different between subpopulations, do not drive my results.

The first encouraging fact is that the sample is well balanced. The number of observations in each of the subgroups (treatment and control for both concentrated and competitive banking markets) is statistically identical (see Table 1). This is

not a result of the sampling process, as the variables that determine the subgroups (geographical area of competition and bank concentration) are not related to those used to determine the sampling (number of employees, urban or rural county and census division). Furthermore, the fact that the panel remains well balanced when only those in the subpopulation of interest are considered means that there is not likely a self-selection process that may be driving the results. Finally, if we look at Table 7, the fact that risk ratings are the same across subpopulations (both amongst all firms and those that received financing) argues against systematically different risk profiles between treatment and control groups; whereas the fact that firm age and growth rates are also indistinguishable across subpopulations (see Table 2, Panel A) should allay fears about systematic differences in investment opportunities as well.

However, not all observable characteristics are similar across groups. To ascertain whether this constitutes a problem for the robustness of my results I perform non-parametric tests on the effect those variables that vary between groups may have on equilibrium credit spreads.

Table 2 panel A shows that an overwhelming majority of observations in rural areas are also in concentrated banking markets. Although we control for urban/rural status in all specifications of our test, one might worry that the parametric control is not enough and that some of my results might be driven by the rural nature of the local firms in concentrated banking markets. In order to address this concern, the second column in Table 4 shows that the results of fitting the baseline model to a restricted sample consisting exclusively of observations in urban locations continue to support the hypothesis of bank "overcharging" local firms, with the coefficient

Table 4: Non-parametric controls for firm characteristics

			Firm Assets	
Models	Full	Urban	Large	Small
Main effect of concentrated	1.33	2.09*	1.73**	1.39
banking market	(2.09)	(2.44)	(4.69)	(2.31)
Does the firm compete	−0.35*	−0.38**	0.02	−0.47*
in local markets?	(−2.51)	(−3.44)	(0.13)	(−2.91)
Interaction effect	0.69**	1.23**	0.62***	0.95**
	(4.14)	(4.15)	(6.06)	(3.96)
Observations	1311	1012	661	650
R^2	0.561	0.581	0.576	0.597

t statistics in parentheses. Significance denoted by * $p < .1$, ** $p < .05$, *** $p < .01$

of the interaction term entering the specification at almost twice the level of the unrestricted sample (123 basis points) and statistically significant at the 1 percent level.

Next, I analyze the most important variable where the four subpopulations of interest in my study are indeed significantly different, pointing at potentially different data generating processes: firm size (see Table 2). This is a critical variable in the corporate finance literature as it can proxy for a number of key structural parameters, such as information opacity, bargaining power or probability of success. In order to test whether it is variation in firm size that drives my empirical results, I control non-parametrically for firm assets. The third and fourth columns on Table 4 show how the coefficients for the main and interaction variables still exhibit the same pattern when I break the population into two groups based on median assets and estimate two separate regressions, in effect interacting this categorical version of the size variable with all other variables in my specification. The economic and statistical

significance of the effect persists on both subsamples, with the smaller firms being subject to larger distortions from bank concentration: a coefficient on the interaction term of 95 basis points, as opposed to 62 for larger firms and 69 for the entire sample.

It can be argued that different loan types follow different pricing systems and thus the fact that lines of credit are significantly less prevalent amongst local firms in concentrated banking markets (representing 49 percent of the loan population) than in all other subgroups (where they account for between 60 and 70 percent of loans, see Panel C of Table 2) could point to an omitted interaction that could be driving my results. The first three columns of Table 5 summarize the results of a non-parametrical test of this alternative hypothesis. Columns two and three show the coefficients for the main and interaction variables when the model is fitted independently for lines of credit and all other loan types. Even though the effects of bank concentration on local firms are almost four times bigger for term loans than for lines of credit (120 basis points versus 29), in both subgroups the coefficient on the interaction term is positive and statistically significant at the five percent level, confirming that even if different pricing systems are in effect they do not drive the main results of this paper.

Finally, in columns four and five of Table 5 I perform the same analysis for long term (more than twelve months) and short term loans. Panel C of Table 2 shows that only 49 percent of loans for non-local firms located in competitive banking markets are long term (with maturities longer than 12 months), whereas in all other subgroups this figure lies between 59 and 62 percent. This systematic difference between subgroups could be cause for concern and point towards differences in data

Table 5: Non-parametric controls for loan characteristics

		Type of loan		Maturity	
Models	Full	Line of Credit	Other	Short	Long
Main effect concentrated	1.33	1.30	0.35	2.31**	0.94
banking market	(2.09)	(1.35)	(0.66)	(3.94)	(1.55)
Does the firm compete	−0.35*	−0.35**	−0.48	−0.84*	0.02
in local markets?	(−2.51)	(−4.16)	(−1.57)	(−2.36)	(0.14)
Interaction effect	0.69**	0.29**	1.20**	1.27**	0.46*
	(4.14)	(5.02)	(3.26)	(5.82)	(3.12)
Observations	1311	850	461	622	689
R^2	0.561	0.614	0.647	0.550	0.672

t statistics in parentheses. Significance denoted by * $p < .1$, ** $p < .05$, *** $p < .01$

generating processes that explain my results without recourse to the exercise of market power. However, the non-parametrical test reported in the last two columns of Table 5 shows that, even though the coefficients on the interaction term suggest that short term loans are more profoundly affected by the exercise of market power (127 basis points as opposed to only 46 for long term loans), both coefficients remain positive and significant (at the five percent level for short term loans and slightly above it for long term loans).

4.2 Alternative explanations

I interpret the results of the previous sections as providing evidence that bank concentration leads to the exercise of significant market power in the pricing of loans to firms whose geographical footprint coincides with that of the oligopolistic banks. However, the main identifying assumption of the empirical analysis cannot be itself tested and, if not true, there are alternative explanations for the test results. Below

I present these alternative explanations, along with the reasons I believe they are less likely to account for the results of this paper.

There are two main kinds of alternative hypotheses: those that posit that there is no (significant) exercise of market power (i.e. that the commercial loan market for small firms is competitive) and those that posit that indeed market power is being exercised by concentrated banks, but that some kind of friction (and not industrial market structure) explains the higher interest rates that local firms experience.

I will deal with the second competing hypothesis first. If we subscribe to the view that bank concentration does indeed have an effect on small firm lending and concentrated banks are able to extract surplus from their borrowers, this competing hypothesis essentially posits that the binding constraint on the interest rate that concentrated banks can charge its borrowers is not determined by the level of industrial competition, but by the limit pricing that forestalls competition from other banks. In order to explain the differential effect on firms that sell their products only locally that this paper measures, it must be the case that these firms face a higher level of whatever barrier to entry is the original source of the bank's monopoly power: most likely, information opacity. If a firm's demand is exclusively located in a local market the cost for potential competing banks from any other market of becoming informed are so much higher, thus constituting a source of monopoly power that could explain my results without recourse to industrial competition: the key to bank monopoly power is information opacity and firms that sell their products in wide areas are not as opaque as local firms and therefore suffer less rent extraction from banks.

The first problem that this hypothesis faces is that it provides no explanation for

why there is any surplus at all to be expropriated in the first place. In order to find an economically significant rent being extracted by concentrated banks, proponents of this explanation have to make ad hoc assumptions about the level of industrial competition in the US; whereas my explanation of the empirical results provides a reason both for the existence of industrial surplus and the banks' ability to extract it.

To further test the difference between this competing explanation of the data and the one proposed in this paper I need to find a group of observations with high informational opacity that are nevertheless not being systematically overcharged by concentrated banks due to high industrial competition. Column two of Table 6 shows that if we restrict our attention to the young firms in our sample (those at or below the median firm age of 16 years) that we would expect to suffer from large information opacity, the interaction coefficient is still positive and significant (121 basis points compared to 69 for the full sample) whereas the main coefficient for bank concentration is not significant. This would lead us to the conclusion that even in this case of information opacity, non-local firms are not being overcharged by their locally concentrated banks. This effect persists if we look at more extreme cases of information opacity due to firm age. Column three of Table 6 shows the same result for firms with 9 or less than 9 years of age (25 percentile) with similar, if less statistically significant results due to the smaller sample size.

If we adopt the position that the small firm credit market is reasonably competitive (the first alternative hypothesis), it would be the case that no monopoly power is effectively exercised by banks and the variability in interest rates charged corre-

Table 6: Other non-parametric controls

		Part of a Bank Holding Co		Firm Age	
Models	Full	Yes	No	$\leq$ 16 yrs	$\leq$ 9 yrs
Main effect concentrated	1.33	0.91*	1.11	0.39	−0.04
banking market	(2.09)	(2.37)	(0.57)	(0.96)	(−0.04)
Does the firm compete	−0.35*	−0.03	−2.02*	−0.54	−0.27
in local markets?	(−2.51)	(−0.20)	(−2.78)	(−1.91)	(−0.66)
Interaction effect	0.69**	0.39**	3.33***	1.21**	1.18*
	(4.14)	(3.20)	(8.63)	(3.55)	(2.54)
Observations	1311	1054	216	654	350
R^2	0.561	0.607	0.729	0.666	0.750

t statistics in parentheses. Significance denoted by * $p < .1$, ** $p < .05$, *** $p < .01$

sponds to observable and unobservable differences in the equilibrium determinants of competitive interest rates: bank costs, price and availability of alternative sources of funding, firm risk profiles and investment opportunities. I will examine each of these possibilities in turn.

First we need to consider that, in order to explain the results of this paper, an unobserved variable would require three characteristics: First, it would need not to have been controlled for by all other sources of risk and relationship already accounted for; second, it would need to justify higher rates for firms that compete in local markets and not for those that compete in broader markets; finally, it would need to affect only those firms that operate in locations where the banking sector is most concentrated. A good candidate would be an unobserved variable that increases the risk of local businesses or reduces their available investment opportunities and, at the same time, restricts the equilibrium number of banks that the local economy can support, so that the end result is that this unobserved variable influences both

the level of bank concentration and the risk or potential reward associated with the earnings of firms that compete in local markets and therefore their competitive cost of funding.

Bank cost structure could be related to both bank concentration and higher interest rates. Small markets (or those with low saving rates) naturally support few banks and could mean high funding costs or operating costs (if fixed establishment costs have to spread over low loan volumes) that lead to higher interest rates without recourse to the exercise of market power. However, this hypothesis does not explain differences between local and non-local firms as none of the higher costs are specific to any firm type. Costs that could be directly associated to a particular firm (information gathering, for example) are likely to be lower for firms that compete in local markets than for those whose competitors span a wider area, suggesting that, if anything heterogeneity in banking costs should lead to *lower* interest rates for local firms, not higher.

Another alternative explanation for the results in this paper is an increase in credit demand due to a shift *away* from alternative funding sources. We can consider that firms in our population of interest have minimal access to organized markets for either debt or equity[12]. SSBF data allows us to answer this concern directly, as firms are asked about their demand for debt financing. The survey contains data on whether companies applied for any loan in the past three years or even if they did not apply through fear of denial but would have wanted to. Panel B of Table

[12]Although we have no data on bond issuance, out of the 4,231 firms in our sample, only 9 were traded in organized equity markets and due to their size it is reasonable to assume organized bond markets are also closed to them, specially to those that compete only at the local level.

2 shows that the percentage of firms in each subpopulation that demanded loans is very similar, casting doubts over this alternative hypothesis.

Yet another alternative explanation of the empirical results of this paper is that there is a systematic difference in the amount or quality of investment opportunities for local firms in concentrated banking areas that explains the difference in interest rates without the need to recourse to banking market power. This explanation faces two problems. The first is that traditional proxies for investment opportunities (such as historical growth rates and average firm age) are statistically identical across subpopulations (see panel A of Table 1). The second and more serious problem is that, in order to justify higher interest rates, investment opportunities for local firms should be higher than for other comparable firms leading to increased demand for funds (also not supported by the data, as discussed above), a situation difficult to reconcile with locations with a small number of banks in equilibrium.

The last alternative explanation for the test results that claims banking markets are competitive is that firms that compete in local markets where banks are concentrated are somehow riskier than those that compete in local but competitive banking markets and therefore their higher interest rates are justified. In order to ascertain the relative merit of this explanation we need to dig deeper into the characteristics that this unique risk factor(s) should have.

In order for this unobservable risk source to affect only firms that compete locally, it must come from the revenue side (as cost factors are likely to affect all firms irrespectively of where do they sell their products). In this case, we would expect some of this systematic difference in risk between subpopulations to be reflected in

objective risk measures that the model controls for, like the the Dun and Bradstreet credit score. And yet, as Table 7 shows, the distribution of risk categories between subpopulations does not support the hypothesis that firms that compete locally in concentrated bank markets are riskier, neither on average, nor notch by notch. If anything, it appears that they are less risky than their homologues located in competitive banking markets, as they show a statistically significant higher proportion of top rated firms. This is true *both* in the entire population of firms and in those that are bank financed. Finally, even if the unobservable risk source were not controlled for by the credit score analysis, one would expect that the rejection rate of firms in the riskier subpopulation would be higher. Yet, panel B of Table 2 shows that there is no statistically significant difference in rejection rates between local firms based on bank concentration, in fact the rejection rate of firms with wide geographical footprints located in areas with competitive banking markets is the largest of all the subgroups. More broadly, there appear to be very few other differences between the distributions of firms in each subpopulation *whether or not they get bank finance* (see Table 2).

Yet, the risk factor that makes local firms riskier may just be the fact that they have such a concentrated market. The argument goes like this: banks that lend to local firms bear more risk because their portfolio is not diversified with respect to local economic shocks therefore they should charge higher interest rates. There are two problems with this argument. First, if this was the main driver of higher interest rates it would apply to regions with concentrated *and* competitive banking markets. But it does not, in fact firms that sell their products locally are charged

systematically lower rates than those with a wide geographical footprint (see the coefficients for the local main effect on Table 3). Second, loans originated by bank holding companies (those with many branches) and therefore likely to be able to diversify away whatever unobservable "local" risk factor might have been driving interest rates up, still exhibit the same results as reported above.

Columns two and three of Table 6 show that the coefficient of the interaction term when the model is fitted independently for loans with banks that belong to bank holding companies and those that do not remains positive and statistically significant at the five percent level or better for both populations. The effect on loans with banks that do not belong to bank holding companies is much larger than the one reported in the main section of this paper (333 basis points as opposed to 69), showing that indeed they consider local loans riskier, but even for bank holding companies the effect of bank market power is significant (39 basis points).

5 Conclusion

My paper contributes to the debate over the appropriate banking market definition by drawing attention to the insights of the vertical relationship paradigm of industrial organization. Even if geographically concentrated banks are able to establish barriers that insulate them from competition from outsiders and even from the threat of new entrants, they might still not be able to overcharge their small firm borrowers. The reason is that competing banks, even if they are unable to overcome the geographical barriers to lending, might be able to lend to a firm's competitors

Table 7: Distribution of Credit Ratings

	Concentrated banking market				Competitive banking market			
	Local firm		Not local		Local firm		Not local	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Average D&B credit score								
All respondents	2.79	0.04	2.97	0.05	2.85	0.04	2.77	0.05
Recently approved loans	2.79	0.09	3.09	0.08	2.93	0.08	2.91	0.10
Percentage of population with safest credit score = 1								
All respondents	7.8	1.2	6.4	1.0	8.6	1.2	11.6	1.5
Recently approved loans	11.7	2.8	4.5	1.4	5.2	1.5	13.1	3.0
Percentage of population with credit score = 2								
All respondents	40.3	2.0	35.0	2.0	37.3	1.9	34.5	2.0
Recently approved loans	37.9	4.0	33.3	3.7	38.6	3.9	27.6	3.1
Percentage of population with credit score = 3								
All respondents	25.1	1.8	24.1	1.7	24.3	1.6	25.9	1.8
Recently approved loans	22.3	3.3	25.1	3.0	24.8	3.0	27.3	3.2
Percentage of population with credit score = 4								
All respondents	14.7	1.4	21.9	1.7	18.4	1.4	16.6	1.4
Recently approved loans	15.4	2.6	22.4	3.2	19.2	2.7	17.5	2.5
Percentage of population with riskiest credit score = 5								
All respondents	10.4	1.1	11.5	1.2	10.8	1.1	9.4	1.2
Recently approved loans	12.3	2.5	14.6	2.5	11.3	2.1	13.8	2.8

at lower rates than the concentrated bank, thus conferring these firms a strategic advantage in the industrial competitive game and effectively limiting the available surplus and therefore the upper limit of the what the concentrated banks can charge. Therefore, the *maximum* amount of overcharging that can be successfully exercised by a concentrated bank is limited by the *minimum* of the market structures in the financial and the industrial markets.

Using data on firm, loan and bank characteristics from the Survey of Small Business Finances I investigate if there are systematic differences in lending conditions between firms whose competitive footprint falls under concentrated or competitive banking markets that cannot be attributed to the usual determinants of lending. I focus on these two subsets of firms because they are those most likely to face significant and sustainable overpricing by banks *if* they are able to exercise monopoly power.

I find empirical evidence to support the claim that concentrated banks do indeed successfully exercise their market power to extract surplus away from local borrowers. But they can only do this as long as the borrowers themselves compete mainly with other local businesses, also under the influence of the concentrated banks.

One implication of this finding is that, when analyzing antitrust issues in banking a new measure of bank concentration is needed: one that analyzes loan market sizes in terms of industries (i.e. competitive space of borrowers) and not geographical footprints. It is only over entire industries that concentration of financiers becomes problematic.

In terms of antitrust policy, this paper shows, both conceptually and empirically

the difficulty in sustaining concentration in the financial markets, without some related kind of concentration present in industrial markets. It is clear from the definition of economic market for commercial lending presented here that a little financial competition can spread very quickly amongst geographically interconnected and competitive industries. And the results of our empirical tests support the hypothesis that there is no significant pricing difference in loans to firms in interconnected and/or competitive markets regardless of the bank concentration they are exposed to.

A second implication is that the best strategy for concentrated local financiers is to enforce collusion and market sharing amongst their local borrowers and expropriate most of the surplus generated, thus hurting consumers as they will perpetuate monopolistic pricing in the local downstream industries. Although this paper finds evidence for overcharging of local firms by concentrated banks, only conceptual conjectures can be made as to the effects on local industrial markets. Further research that could directly test whether in fact small firms that compete locally in areas where banks are concentrated charge prices closer to the collusive equilibrium than to the competitive level could be illuminating.

The second main area open for further research is the precise nature of the barrier to competition from out-of-town banks that allows overcharging. The proposed conceptual framework leaves a lot of uncertainty about the relative importance of many relevant theoretical effects that can determine the practical importance of the interrelation between concentration in the product and financial industries. In particular, it would be interesting to further analyze whether issues of firm quality or character (and therefore relationship banking) or issues of industry risk analysis (and

therefore bank specialization) have a larger impact on interest rates.

References

R. M. Adams, L.-H. Röller, and R. C. Sickles. Market power in outputs and inputs: An empirical application to banking. FEDS Working Paper No. 2002-52, October 2002.

D. F. Amel and M. Starr-McCluer. Market definition in banking: Recent evidence. *The Antitrust Bulletin*, 47:63, 2002.

J. S. Bain. A note on pricing in monopoly and oligopoly. *American Economic Review*, 39(2):448–464, 1949.

J. S. Bain. Relation of profit rate to industry concentration. *Quarterly Journal of Economics*, 65:239–324, August 1951.

W. J. Baumol, J. C. Panzar, and R. D. Willig. *Contestable markets and the theory of market structure*. New York: Harcourt Brace Javanovich, Inc, 1982.

A. N. Berger and T. H. Hannan. The price concentration relationship in banking. *Review of Economics and Statistics*, 21, 1989.

A. N. Berger and G. F. Udell. Relationship lending and lines of credit in small firm finance. *Journal of Business*, 68(3):351–382, 1995.

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

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

A. W. A. Boot. Relationship banking: What do we know? *Journal of Financial Intermediation*, 9:7–25, 2000.

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

T. F. Bresnahan. The oligopoly solution concept is identified. *Economics Letters*, 10:87–92, 1982.

N. Cetorelli and P. E. Strahan. Finance as a barrier to entry: Bank competition and industry structure in local US markets. *The Journal of Finance*, 61(1):437–461, 2006.

R. A. Cole, J. D. Wolken, and R. L. Woodburn. Bank and nonbank competition for small business credit: Evidence from the 1987 and 1993 national surveys of small business finances. *Federal Reserve Bulletin*, 82(11):983–995, Novemeber 1996.

S. G. Craig and P. Hardee. The impact of bank consolidation on small business credit availability. *Journal of Banking & Finance*, 31(4):1237–1263, April 2007.

H. Degryse and S. Ongena. Distance, lending relationships and competition. *The Journal of Finance*, 60(1):231–266, February 2005.

H. Degryse and P. Van Cayseele. Relationship lending within a bank-based system: Evidence from European small business data. *Journal of Financial Intermediation*, 9(1):90–109, January 2000.

G. Dell'Ariccia. Asymmetric information and the structure of the banking industry. *European Economic Review*, 45(10):1957–1980, 2001.

G. Dell'Ariccia, E. Friedman, and R. Marquez. Adverse selection as a barrier to entry in the banking industry. *The RAND Journal of Economics*, 30(3):515–534, 1999.

R. Gilbert. Bank market structure and competition: A survey. *Journal of Money, Credit and Banking*, 16(4):617–645, 1984.

G. Hale and J. Santos. Do banks price their informational monopoly? *Journal of Financial Economics*, 93(2):185–206, 2009.

T. Hannan. Bank commercial loan markets and the role of market structure: Evidence from surveys of commercial lending. *Journal of Banking & Finance*, 15(1): 133–149, 1991.

R. Hauswald and R. Marquez. Competition and strategic information acquisition in credit markets. *Review of Financial Studies*, 19(3):967–1000, 2006.

W. E. Jackson. Is the market well defind in bank merger and acquisition analysis? *The Review of Economics and Statistics*, 74(4):655–661, 1992.

J. Jayaratne and P. E. Strahan. The finance-growth nexus: Evidence from bank branch deregulation. *Quarterly Journal of Economics*, 111:639–670, 1996.

W. Kerr and R. Nanda. Democratizing entry: Banking deregulation, financing constraints, and entrepreneurship. *Journal of Financial Economics*, 94(124–149), 2009.

M. Klein. A theory of the banking firm. *Journal of Money, Credit and Banking*, 3: 205–218, 1971.

R. Marquez. Competition, adverse selection and information dispersion in the banking industry. *Review of Financial Studies*, 15(3):901–926, 2002.

E. S. Mason. Price and production policies of large-scale enterprise. *American Economic Review*, 29:61–74, 1939.

B. D. Meyer. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13(2):151–161, April 1995.

M. Monti. *Mathematical methods in investment and finance*, chapter Deposit, credit and interest rate determination under alternative bank objective functions, pages 431–454. Amsterdam: North Holland, 1972.

M. Petersen and R. G. Rajan. The benefits of lending relationships: Evidence from small business data. *Journal of Finance*, 49(1):3–37, 1994.

M. Petersen and R. G. Rajan. The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443, 1995.

R. G. Rajan. Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of finance*, pages 1367–1400, 1992.

R. G. Rajan and L. Zingales. Financial dependence and growth. *American Economic Review*, 88(3):559–586, 1998.

T. Rice and P. E. Strahan. Does credit competition affect small-firm finance? *The Journal of Finance*, 65(3):861–889, 2010.

J. N. Rosse and J. C. Panzar. Chamberlain vs. Robinson: An empirical study for monopoly rents. Bell Laboratories Economic Discussion Paper, 1977.

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

R. Sengupta. Foreign entry and bank competition. *Journal of Financial Economics*, 84(2):502–528, 2007.

S. Shaffer. Patterns of competition in banking. *Journal of Economics & Business*, 56:287–313, 2004.

G. J. Stigler and R. A. Sherwin. The extent of the market. *The Journal of Law and Economics*, 28:555–585, October 1985.

J. Tirole. *The theory of industrial organization*. Cambridge, MA: MIT Press, 1988.

E. Van Tassel. Relationship lending under asymmetric information: a case of blockaded entry. *International Journal of Industrial Organization*, 24:915–929, 2006.

L. W. Weiss. *Concentration and price*, chapter A review of concentration-price studies in banking, pages 219–254. Cambridge, MA: MIT Press, 1989.

R. Zarutskie. Evidence on the effects of bank competition on firm borrowing and investment. *Journal of Financial Economics*, 81(3):503–537, 2006.

A Appendix

In this appendix I present a simple model that formalizes the definition of banking market used throughout the paper.

Let us assume a risk-neutral world. There are two types of agents in this economy: firms that produce goods and financial intermediaries that provide funding. For ease of exposition I will assume that there are only a small number of places where agents can be located, we can think of them as cities or towns. I will further assume that each agent acts out of one, and only one location.

There are two types of goods in this economy. Some goods (let us call them global goods) can be transported and can therefore be consumed at any location (regardless of where they are produced) by incurring certain transportation costs. There are other goods however, that can only be consumed in the location where they are produced. We can think of any real world commodity (such as corn, crude oil, etc.) as being an example of a "global good", whereas "local goods" could be represented by, for instance, the services provided by restaurants or coffee shops. Exogeneously determined consumer preferences determine demand for each good at each location. This demand is however uncertain. For simplicity, I will assume that some sufficiently high level of demand occurs for the goods that firm j sells domestically with probability q_j. With probability $1 - q_j$ however, there will be no demand for firm j's goods.

A.1 Firms

Firms, indexed by j in this model, produce only one type of good out of only one location with the objective of maximizing expected profits. The location of firms is exogenously determined and may depend on factors such as expected demand or technology. Firms are also assumed to be owner-managed. In this way, the residual claimants for the firm's income are the ones making both the investment and financing decisions and I can abstract away from a number of agency issues.

In order to produce and sell its goods, firm j has to incur capital expenditures K_j. I will assume that firms do not have enough internal funds to finance these outlays and must therefore tap external sources of finance. I will also assume that the only source of external finance available is the banking industry.

Once investment has been completed, firms make their operating decisions concerning production and pricing, and finally uncertainty surrounding local demand is resolved and returns are determined.

Firm's j potential gain from production depends not only on the level of demand, but also on what other competing firms "within range" are doing (for example, firm k); where the definition of what can be considered "within range" depends on the magnitude of transportation costs between locations: $t(j,k)$. I will assume that firms compete on prices. Therefore, firm prices (and revenues) are capped by those charged by a firm's competitors, after taking into account the effects of distance.

$$p_j \leq p_k + t(j,k), \text{ where } j \neq k \tag{1}$$

I will assume in this model that operating costs, efficiency levels, capital intensity and outside options for the entrepreneurs are identical across firms that produce the same goods.

What a firm can charge is also capped by the financing it gets. There is a lower limit on prices determined by the need to make a non-negative expected profit after financing. Financing conditions, D_{ij}, are determined for each pairing of firm j and bank i by means of a bargaining game.

Therefore, the net profit that accrues to the firm is described by the following equation, where c_j, x_j and p_j represent, respectively, costs, production and pricing choices, and the firm is protected by limited liability.

$$\Pi_{ij}^{f} = \begin{cases} (p_j - c_j)x_j - D_{ij} & \text{with prob } q_j \\ 0 & \text{with prob } 1 - q_j \end{cases} \tag{2}$$

In order to maximize expected profit, the entrepreneur first bargains over loan terms with bank i, agreeing upon rate of interest of r_{ij} for a size K_j loan, and chooses a pricing policy, p_j. In principle, the entrepreneur could agree to any loan rate that yields her expected profits in excess of her best outside option, A_j. This determines the entrepreneur's participation constraint (PC_f). The actual rate agreed upon will also depend on the bank's participation constraint and the relative bargaining power of both parties.

$$q_j(p_j - c_j)x_j - q_j D_{ij} \geq A_j \tag{3}$$

A.2 Banks

For simplicity, I will refer to the financial intermediaries in this economy collectively as banks. The objectives and constraints of the bank are standard: the bank wants to maximize its expected profits from loans (again, assuming risk neutrality). The location of banks is exogenous in this model and I will assume that each bank operates out of one and only one location. However, banks can lend to firms in any location by incurring some extra costs that will be discussed below. In one extreme case, banks across a certain boundary would not be able to compete for the opportunity to lend to firms on the other side of that boundary, maybe due to regulatory barriers for example. The alternative extreme is when bank lending costs are identical regardless of distance to borrower, in that case we can consider all financial markets interconnected and there is very little cause for worrying about bank concentration in such a setting.

The bank's expected revenue is the rate on the loan r_{ij}, or the salvage value of the firm in the case of bankruptcy, s_j. On the cost side, banks themselves incur financing costs, and must also cover their own operating expenses.

For the purpose of this model, it is enough to distinguish between bank operating costs that vary with distance between borrower and bank, $d(i,j)\delta$ (directly or indirectly, maybe because of lack of local information) and those that do not, b_{ij}. Most of a bank's explicit costs (those easier to measure) such as funding and operating costs are unlikely to depend on distance. But there are several important cost sources that can vary with distance to borrower. Both relationship generation costs and monitoring costs should increase with distance, as they involve face to

face meetings and similar hands-on interaction. Information acquisition costs can also increase with distance. Finally, a large value of distance related costs could be determined by search processes, if we posit that customers do not routinely look for finance in a very large area, and banks find it expensive to "solicit" for unrequested business at long distances.

The profit, Π^b_{ij}, bank i will derive from lending to firm j depends on the amount of the loan (K_j), the interest rate charged (r_{ij}), the salvage value of the firm in case of bankruptcy (s_j), the bank's fixed operating costs (b_{ij}) and those related to distance ($d(i,j)\delta$), and finally, the probability of success (q_j) as described by:

$$\Pi^b_{ij} = \begin{cases} K_j(r_{ij} - b_{ij} - d(i,j)\delta) & \text{with prob } q_j \\ K_j(sj - b_{ij} - d(i,j)\delta) & \text{with prob } 1 - q_j \end{cases} \tag{4}$$

In order to break even in expectation, bank i needs to be promised returns ($D_{ij} \equiv r_{ij}K_j$) in the amount of:

$$\underline{\mathrm{D}}_{ij} = K_j\Big(\frac{c_{ij}}{q_j} - \frac{(1-q_j)}{q_j}s_j\Big) = K_j\Big(\frac{c_{ij}}{q_j} - \alpha_j s_j\Big) \tag{5}$$

A.3 Characterizing the equilibrium rates

This model describes a situation in which entrepreneurs and financiers bargain with each other in order to secure funds for investment into competitive marketplaces. Each firm-bank pair constitutes a bargaining game. Banks also play a competitive game amongst each other as do firms.

The equilibrium level of interest rates that we can expect to observe as a result

of these interrelated games can only be determined empirically, as it would be determined by the precise bargaining mechanism in each type of game and the distribution of bargaining power between players. However, analyzing these interrelated games can refine our understanding of the bounds on the feasible solutions. In this framework there are only two types of games being played: one by banks and each firm about financing and its terms, and another between those firms that get financed about quantities and prices. The goal of my modeling is to isolate the key variables that determine the exercise of market power to design an appropriate empirical test.

In order to characterize more precisely the limits to the equilibrium behavior of players in my model I will introduce a number of simplifying assumptions. I will consider only two locations (A and B), each one with only one firm producing a certain "global good". This does not mean, however that firms are local monopolists.

In order to characterize the differences between concentrated and competitive banking markets I will assume that in location A there is only one bank, whereas location B has many banks (see Figure 1). To sharply focus on the effects of the interaction between competition in industrial and financial markets, I will make very strong assumptions about how players are different from one another. In particular, I assume that capital requirements are identical for both firms (all K_j are the same), operating costs are the same for both industrial firms and all outside options are zero. This means that the expected profit of firm j becomes $q_j(R_j - D_{ij})$.

I also assume that bank operating costs are identical (all K_i and δ_i are the same), that all loans are uncollateralized, that there is no credit rationing in the financial market and that banks have access to an unlimited pool of capital as long as they

can earn a non-zero expected net return on it.

In order to characterize the limits to equilibrium interest rates, I will first calculate the minimum unlevered revenue that firm 2, located in the competitive banking area, must earn in the good state of the world, in order to break even. It is important to realize that this quantity embodies the appropriate competitive risk-adjusted price of capital.

$$\underline{\mathrm{R}}_2 = \frac{b}{q_2} + \frac{\delta}{q_2} d(B, 2) \tag{6}$$

If we assume that distance within the same location is negligible the result is as expected: in a fully competitive equilibrium firms must earn enough to cover their own and their banks operating costs in expectation.

I then calculate the loan rate that competitive banks (located in B) would have to charge firms in location A (where there is a financial monopoly) in order to break even, despite the distance. This is the concept of limit pricing (as first introduced by Bain (1949)). The idea is that the incumbent, in our case the monopolist bank, will charge its clients this rate: the highest price that cannot be matched by competitor financiers; be it currently established competitors or potential new entrants into the financing market.

$$\bar{\bar{D}}_{A1} = \frac{b}{q_1} + \frac{\delta}{q_1} d(B, 1) \tag{7}$$

Finally, I calculate the maximum rate that the monopolist bank can charge its local firm (firm 1) in the absence of financial competitors, but subject to industrial

competition. There are many ways to model the nuances of competition in the industrial market. For our purposes, the only thing that matters is that the unlevered profits (before financing) of firm j react positively to increases in the interest rates charged to its competitors.

As long as this condition is satisfied, the top of the bargaining range depends on the level of competition in the banking market, through the minimum rate that other banks can charge. In our setup, this is determined by firm 1's available surplus, which is restricted to be at most equal to that of firm 2, its potential competitor financed competitively, plus transportation costs.

$$\bar{D}_{A1} = \bar{R}_1 - A = \frac{B}{q_2} + \frac{\delta}{q_2} d(B,2) + t(1,2) + \epsilon_2 \tag{8}$$

The upper bound to equilibrium interest rates is thus defined by the *minimum* of the maximum rate local borrowers would agree to ($\bar{D}_{A1}$) *and* the limit price that will keep "foreign" banks away ($\bar{\bar{D}}_{A1}$). Therefore, depending upon which of the two upper bounds on interest rates is binding, the scope for a monopolist bank to overcharge its local firms is defined by:

$$\bar{\bar{M}}_A = \frac{\delta}{q_1}\big(d(B,1) - d(A,1)\big), \tag{9}$$

if bank competition is the binding constraint, or by:

$$\bar{M}_A = B\big(q_2^{-1} - q_1^{-1}\big) + t(1,2) + \epsilon_2 \tag{10}$$

if the binding constraint is industrial competition and we further assume $d(A,1) =$

$d(B, 2) = 0$, where ϵ_2 represents the industrial market power of firms under competitive financing.

The rates observed in the real world will be bound by this restriction, but will depend on the distribution of bargaining power amongst observed firm-banks pairs. It is not immediately apparent which of these two restrictions should be binding. The answer will hinge on how variable is the level of competition in the financial sector compared to the level of competition in the industrial sectors. Whether the size of the variation in bargaining power is "large" relative to the variation in pre-existing economic surplus or the opposite is true is a question that can only be answered empirically.