

Distance and Private Information in Lending

Sumit Agarwal

Federal Reserve Bank of Chicago

Robert Hauswald

American University

We study the effects of physical distance on the acquisition and use of private information in informationally opaque credit markets. Using a unique data set of all loan applications by small firms to a large bank, we show that borrower proximity facilitates the collection of soft information, leading to a trade-off in the availability and pricing of credit, which is more readily accessible to nearby firms albeit at higher interest rates *ceteris paribus*. Analyzing loan rates and firms' decision to switch lenders provides further evidence for banks' strategic use of private information. However, distance erodes our lender's ability to collect proprietary intelligence and to carve out local captive markets, suggesting that the requisite soft information is primarily local. (*JEL* G21, L11, L14, D44)

Private information and its distribution are among the fundamental forces shaping economic exchange. Agents often devote considerable resources in terms of time, effort, and money to its acquisition in order to gain a strategic advantage in the ensuing transaction, especially by collecting soft information. At the same time, very little is known about the origins, use, and consequences of such subjective intelligence despite its economic importance (Aghion and Tirole 1997; Stein 2002), in part because it has eluded easy categorization so far (but see Petersen 2004). Its defining attributes—it is not readily transferable, verifiable, or interpretable—also imply that it is difficult to identify, measure, and analyze in practice. However, technological progress coupled with operational procedures in commercial lending allow us to overcome these analytic challenges in one particular industry—credit-market transactions in-

We thank Hans Degryse, Thomas Gehrig, Dennis Glennon, Uli Hege, Victoria Ivashina, Esa Jokivuolle, Yan Lee, Bob Losey, Robert Marquez, Kjell Nyborg, Steven Ongena, Charlotte Ostergaard, Marco Pagano, Greg Udell, and Bas Werker, as well as seminar participants at Freiburg, American University, the NHH Bergen, the Central Bank of Norway, the 2006 European Finance Association Meetings, "The Changing Geography of Banking" Conference Ancona, the CREDIT 2006 "Risks in Small Business Lending" Conference, the ECB/CFS Conference "Financial Modernisation and Economic Growth," the 6th FDIC Fall Workshop, the 43rd Conference on Bank Structure & Competition, and the 2008 Western Finance Association Meetings for comments. We are particularly grateful to Mitch Petersen for detailed comments, which improved the exposition and analysis, to the editor, and to an anonymous referee, whose suggestions considerably sharpened the results and their interpretation. Jeff Chin provided outstanding research assistance as did Johann-Alexander and Philippe Hauswald. We gratefully acknowledge financial support from the FDIC's Center for Financial Research. The views expressed in this research are those of the authors and do not necessarily represent the policies or positions of the Federal Reserve Board or the Federal Reserve Bank of Chicago. Send correspondence to Sumit Agarwal, Federal Reserve Bank of Chicago, Chicago, IL 60604-1413; telephone: (312) 322-5973; fax: (312) 322-6003. E-mail: ushakri@yahoo.com.

volving small, informationally opaque firms—and to provide direct evidence on the collection and use of soft information.

Lending to small businesses is particularly well suited for an inquiry into subjective aspects of corporate decision making and their importance for firm policy, investment choice, and, ultimately, performance. On the one hand, banks quantify the information underlying their commercial-lending decisions through credit assessments in the form of credit scores, which allow us to extract meaningful measures of their private information. On the other hand, credit to small businesses crucially relies on firm-specific subjective intelligence collected by loan officers during the origination process and, especially, long-term lending relationships. Proximity to the information source in its various guises should facilitate this endeavor so that we examine the effects of physical distance on the collection of soft private information, its hardening in the credit-assessment process, and its effects on exchange in credit markets.

To shed light on these issues, we rely on a unique data set that consists of all loan applications by small businesses to and credit offers by a major U.S. bank over a fifteen-month period. In addition to the ultimate lending decision and loan terms, we know the results of the bank's internal credit-screening process (proprietary credit score), which comprises a soft component supplied by branch offices through individual score adjustments. To isolate this subjective credit assessment, we first orthogonalize the lender's proprietary score with publicly available estimates of credit quality, which yields a clean measure of its firm-specific private information. By matching our sample with credit-bureau reports, we also identify firms declining the offer and switching lenders to study the market's competitive reaction through a borrower's response to the bank's credit decision. We then estimate discrete-choice models of the bank's and borrower's respective decisions to offer credit and to accept the loan terms and linear-regression models of the offered loan's all-in cost.

We find that distance drives the following fundamental trade-off in the availability and pricing of credit: the closer a firm is to its branch office, the more likely the bank is to offer credit but also the more it charges *ceteris paribus*. Conversely, the closer a firm is to a competitor's branch, the less likely it is to obtain credit but, if it does, the lower the offered loan rate is. Controlling for the firm's public credit quality does not affect these distance effects. However, they disappear to the point of insignificance once we include our measures of proprietary (bank score) and subjective (its orthogonal component) information to control for private intelligence. Furthermore, the loan rate decreases in the bank's subjective (or proprietary) credit assessment of the firm. Interacting distance and information variables shows that a higher subjective or proprietary credit-quality signal reduces the offered rate less for borrowers located close to their branches, which is consistent with lending competition under asymmetric information (e.g., Broecker 1990; von Thadden 2004). Conversely, better private credit assessments influence lending decisions and loan rates less with greater firm–bank distance.

These results suggest the following channel through which distance affects credit decisions: borrower proximity facilitates the collection of soft information, thereby enhancing its quality, which is consistent with the notion that banks enjoy a local informational advantage that distance erodes. To see whether competition correspondingly increases in distance, we next analyze a firm's decision to decline a loan offer and seek credit elsewhere. We find that borrowers located farther away, paying higher rates, or having a higher proprietary score are more likely to switch lenders. Since such customers are the primary targets for rent extraction, these findings imply that the attempt to carve out local captive markets around branches fails at their periphery due to the diminishing quality of their private information, allowing competitors to compete more aggressively and resulting in more switching, as predicted by, for example, [Hauswald and Marquez \(2006\)](#). We interpret our findings as evidence in favor of credit-market models based on locationally differentiated information.

To further assess the importance of location-specific soft information, we next investigate the effect of distance and subjective intelligence on credit delinquency. We find that the farther away a borrower is, the more likely the bank is to experience nonpayment. At the same time, bank–borrower distance greatly reduces the beneficial impact of a higher credit score on predicted delinquency. This finding is consistent with the notion that banks mistrust their private information as distance increases and further points to its local nature.

Our main contribution consists of providing direct evidence on the location-specific nature of soft information, which allows lenders to carve out locally captive markets. We show that customer proximity facilitates the gathering of subjective firm-specific intelligence, which determines the quality of a bank's private information and, thus, its informational advantage. Our findings support the view that soft information in credit markets is primarily local. They also highlight how technological progress in the form of credit scoring allows banks to overcome distance-related limits to lending through the hardening of soft information ([Petersen 2004](#)). To the best of our knowledge, this study is the first one to not only identify the sources of lenders' local information advantage but also empirically show how banks strategically use subjective intelligence. Since lender switching typically does not occur in transportation-cost models, our results also provide strong support for asymmetric-information models of financial intermediation.

Closest to our work is the study by [Petersen and Rajan \(2002\)](#), who first pointed out the local nature of the soft information necessary for credit decisions, and [Degryse and Ongena \(2005\)](#), who show on the basis of Belgian data that distance affects loan transactions through spatial price discrimination. However, neither paper can identify a loan's private-information content and the channel through which it affects bank–borrower interaction. Our results also shed light on the forces that allow banks to increase the geographic extent of their lending activities first noticed by [Petersen and Rajan \(2002\)](#).

We show that these improvements result from the combination of traditional soft-information production through lending relationships with technological advances in credit scoring, which significantly enhance the value of local intelligence for credit decisions. DeYoung, Glennon, and Nigro (2006) report that the probability of default on small-business loans increases in the distance between borrower and lender but that the adoption of credit-scoring techniques reduces this effect, which we also find. By contrast, using Swedish data, Carling and Lundberg (2002) do not find any evidence that distance is a determinant of a loan's default probability.

The effect of physical distance on financial decision making in the presence of local-information advantages has also come to the forefront in other areas. Analyzing informational asymmetries as revealed by proprietary equity trading, Hau (2001) finds that traders located near a company's headquarters outperform their competitors in intraday trading. Similarly, Coval and Moskowitz (1999, 2001) report that fund managers trade local securities at an informational advantage vis-à-vis nonlocal investors who might suffer from asymmetric information. For M&A transactions, Ragozzino and Reuer (2009) report that the higher the perceived threat of adverse selection, the closer acquirers are geographically to their targets, so that the bidder presumably holds an informational advantage over more distant competitors. Garmaise and Moskowitz (2004) also find strong evidence in their analysis of commercial real-estate markets that buyers tend to be local when information asymmetries between the parties are severe and more remote otherwise.

The article is organized as follows. The first section summarizes the role of distance in models of credit-market competition and our testable hypotheses. Section 2 describes our data and estimation strategy. In Section 3, we analyze the determinants of lending decisions and loan pricing. Sections 4 and 5 investigate the determinants of the borrower's decision to accept or reject the bank's loan offer and default behavior, respectively. The last section discusses further implications and concludes. All tables are relegated to the Appendix.

1. The Role of Distance in Loan Transactions

Economic theory has identified two broad channels for physical distance to affect credit-market transactions. As Petersen and Rajan (2002) argue, local lenders can collect "soft" information on small firms over time, thereby permitting them to gain an informational advantage over more remote competitors who might not enjoy the same degree of access to local information.¹ Hauswald and Marquez (2006) make this notion precise by letting the quality of a bank's proprietary information be a decreasing function of the distance between bank and borrower to capture the varying degrees of informational

¹ See Berger, Miller, Petersen, Rajan, and Stein (2005) on the role of soft information in lending decisions and the ability of smaller banks that presumably have a more local focus to collect and process such intelligence.

expertise present in modern banking. Lenders then strategically use their local information advantage to create adverse-selection threats for their rivals, thereby softening price competition and informationally capturing borrowers. As distance erodes the bank's informational advantage, less informed competitors can bid more aggressively (lower interest rates) and more frequently so that, in equilibrium, expected interest rates decrease, and the likelihood of a competing loan offer increases with bank–borrower distance.

The second channel through which distance might affect loan transactions consists of transportation costs incurred by loan applicants and banks. Such costs represent the time and effort spent by a potential borrower to personally interact with loan officers, branch managers, etc., or to look for a suitable loan (product differentiation). Similarly, banks might incur transportation costs in assessing loan applicants or in monitoring loans (Sussman and Zeira 1995), which affect credit terms. Traditionally, such models have focused on uniform loan pricing (e.g., Chiappori, Perez-Castrillo, and Verdier 1995; see Freixas and Rochet 2008 for a survey). But banks invariably know the address of their loan applicants, which allows them to engage in spatial price discrimination (Lederer and Hurter 1986). It is straightforward to show that in this setting, interest rates are also a decreasing function of bank-firm distance (see, e.g., Hauswald and Marquez 2006). However, any borrower deemed creditworthy always obtains credit from the closest bank and would never switch lenders.

Hence, the first empirical difference between the two model classes revolves around the competitive reaction of other lenders as revealed by a loan applicant's decision to switch banks. Since a lender might hold less of an informational advantage at its periphery, competitors can attempt to poach customers more aggressively. In the presence of informational asymmetries, we should therefore observe that borrowers located farther away or of higher credit quality are more likely to decline loan offers. By contrast, if transportation costs mainly drive offered loan terms, distance and proprietary information should not influence an applicant's decision to switch lenders given that the firm must have applied to the optimally located bank in the first place.

The second empirical difference stems from the relation among distance, information, and credit terms that varies by model class. To the extent that borrower proximity facilitates the collection and analysis of local information, distance indirectly affects loan transactions through its impact on the quality of the lender's proprietary intelligence. Hence, the two should be correlated in the presence of informational asymmetries. By the same token, if a lender's informational advantage is primarily local, we would expect banks to make more errors in granting credit the farther away an applicant is so that the likelihood of credit delinquency should also increase with distance. In transportation-cost models, borrower proximity directly affects the terms of a credit offer so that distance should neither be correlated with the bank's information nor affect credit delinquency.

2. Data Description and Methodology

Our sample consists of all applications for new loans over a fifteen-month span falling under the purview of small- and medium-sized enterprise (SME) lending as defined by the Basel I Accord (total obligation and sales below \$1 and 10 million, respectively) to the third largest small-business lender in the United States. Our data provider ranks among the top five commercial banks and savings institutions in terms of deposits during the sample period, operates a total of 1,552 branch offices,² and is the leading provider of SME loans in its area of operation (market shares of 30% to 50% by state). As a result, our lender's business practices, information acquisition, and, hence, loan data are highly representative of the industry, all the more so in that they exemplify the recent success of large institutions in this market segment.³

We focus on small-business lending because small firms tend to be informationally opaque and bound to their local economies. Hence, both information asymmetries and transportation costs might affect loan transactions. Moreover, credit scoring is a particularly successful screening method in this segment of commercial banking. Finally, the product in question—outright loans or credit lines—is highly standardized and, hence, comparable across lenders so that any product differentiation should come from information production, transportation costs, or pricing.

2.1 Lending practices

The application process typically takes four hours to a day to complete from the initial contact between firm and bank. During the branch visit, the manager-owner or firm representative supplies all the relevant data, submits financial and tax information, provides a list of assets, etc., which the local loan officer transcribes into electronic form and matches with credit reports for input into the bank's own proprietary credit-scoring model. Since our lender applies a uniform credit-scoring methodology to assess loan applications, its internal credit score is a meaningful measure of proprietary information across applicants and bank branches.

Concurrently, the loan officer conducts an in-depth interview with the applicant and gathers soft information in the sense that it would be hard to verify by a third party. In about 8% of the cases, the branch office will invite the applicant back to follow up on open questions, review discrepancies in submitted information with credit reports, discuss the prospects of the firm, etc. Such information then allows the branch manager or loan officer to

² For comparability, the one hundred institutions with more than \$10 billion in assets in 2002 operated, on average, 364 branch offices. Their average amount of deposits is about a quarter of our data provider's deposit base.

³ Ely and Robinson (2001) report that banks with assets exceeding \$5 billion held 51% of small-business loans in 1999. By 2003, the largest sixty-seven financial institutions held 64.3% of such loans below \$100,000 (SBA 2005), highlighting the explosive growth and economic importance of large banks in this segment.

subjectively adjust the firm's bank-internal score should the applicant deserve credit in their eyes but fail to meet certain commercial, profitability, liquidity, or credit-score requirements. Each branch enjoys a considerable amount of autonomy in the assessment, approval, and pricing of loans but has to justify any deviation from bank-wide practices on the basis of predefined subjective criteria, such as impression of management quality, personal assessments of collateral value, own view of firm prospects, etc.⁴ Hence, credit decisions ultimately reside with branches but managers' career prospects and remuneration depend on their overall lending success, and local overrides are closely monitored by the bank's risk management.

Day-to-day monitoring is automated through the tracking of current-account movements or balances (whenever available) and prompt debt service. On a monthly basis, the bank collects new credit reports for the firm and its owner and updates the loan's risk profile. On each anniversary of the loan's origination, a firm representative submits updated financial information and discusses the firm's prospects, state of solvency, funding needs, etc., with the account officer in person. If a payment is less than ten days late, the account officer calls the firm to inquire, then personally visits the firm, and eventually cuts back credit lines and calls in the overdue loan.

2.2 Data description

We start with all 28,761 loan applications to our bank that conform to the Basel Accord's SME lending definition and that were made in person at one of its branches from January 2002 to April 2003. Matching these records with Experian and Dunn & Bradstreet credit-bureau reports on the application date leaves approximately 26,028 complete observations.⁵ Next, we verify each applicant's address information with Yahoo! or Google Maps and delete all records with incomplete or nontraceable addresses, leaving a total of 25,746 observations with full data availability.

To identify the nearest bona fide competitor for any loan applicant and to find the driving distances in miles and minutes between the firm, the bank branch, and the nearest competitor's branch, we rely on Yahoo!'s SmartView⁶ and Map services, respectively.⁷ We then look up the driving distance in miles

⁴ Berger, Frame, and Miller (2005) also report that banks using proprietary credit-scoring models often rely on additional discretionary information in the credit approval process.

⁵ Some credit reports were incomplete, did not provide the public credit scores, or contained information that did not match our firm data. Furthermore, our bank engaged in several M&A transactions affecting its branch network so that we also omit all reassigned loan records.

⁶ In contrast to Microsoft's MapPoint, which we also tried, SmartView has the dual advantage that it does not accept sponsored links and draws on the combined yellow-page directories of BellSouth and InfoUSA (Mara 2004), providing objective and comprehensive bank-branch information.

⁷ We only selected full-service branches and, when in doubt, verified smaller institutions' product offerings from their corporate websites, replacing questionable competitors with the next nearest full-service branch of a lender active in small-business lending.

and minutes from the firm to the bank and competitor branches from Yahoo! Maps, which also allows us to calculate the corresponding aerial distances from the geographic coordinates. Driving distances between firms and bank branches range from zero to 3,102 miles, which is clearly too great to conform to standard motives of lending in spatial models. Hence, we remove the top 1% of our sample (259 observations) as a precaution to ensure that our data are compatible with a wide range of locational explanations of credit markets, leaving 25,487 observations with firm–bank distances of up to two hundred fifty miles. Including the outliers does not change our results.

Table 1 provides summary statistics for our main variables as a function of the screening outcome (accept or reject the loan application) and reports the *P*-values of *t*-tests for each variable's mean conditional on the bank's decision. We see that the mean (median) *Firm-Bank Dist* is 9.91 (2.62) miles for successful applicants, as opposed to 10.67 miles for unsuccessful ones, and that this difference is statistically highly significant.⁸ In terms of driving times, loan applicants spend on average about ten minutes (median: seven minutes) and slightly more than two minutes to get to their bank and to the nearest competitor branch, respectively.

The nearest full-service branch of a competitor is an average (median) distance of about 1.10 (0.55) miles away from successful applicants and 0.93 (0.48) miles from unsuccessful ones (*Firm-Competitor Distance*). In terms of driving miles, only 7.21% (driving minutes: 13.17%) of successful loan applicants turn to their closest bank branch as compared to 13.66% of unsuccessful ones (driving minutes: 19.73%); the vast majority prefer to apply for credit from sources farther away. Since transportation-cost models predict that customers choose the closest seller, these statistics imply that small-business loans are not a homogeneous commodity. Instead, other factors must determine loan applications, lending decisions, and, ultimately, price discrimination in our data. By contrast, Degryse and Ongena (2005) report that in their Belgian sample, 25% of borrowers obtain credit from the closest bank, suggesting that in Belgium, transportation costs may play a greater role in loan transactions.

To measure informational effects, we rely on the lender's own creditworthiness assessment summarized by each applicant's final (adjusted) bank-internal credit score, which comprises verifiable ("hard") and subjective ("soft") information. While the scoring methodology is proprietary and subject to confidentiality restrictions, the credit-screening procedure is consistent across all branches and applications, relies on the same approach and statistical model, and uses a common set of inputs except for the subjective element present through local score adjustments.⁹ The *Proprietary Score* ranges from

⁸ For confidentiality reasons, the provider of the data did not permit us to report further descriptive statistics.

⁹ Analyzing periodic surveys of loan officers to assess the importance of subjective (soft) information in credit decisions, our bank estimates that, on average, 20% to 30% of its proprietary scores consist of such intelligence (personal communication).

Table 1
Summary statistics for all loan applications

Loan-application outcome variable	Mean	Accept median	Std Dev	Mean	Reject median	Std Dev	<i>t</i> -Test <i>P</i> -val
Loan Rate (<i>APR</i> : all-in cost of loan)	8.46%	8.12%	2.73%	N/A	N/A	N/A	N/A
Loan Amount	\$46,507	\$39,687	\$42,755	N/A	N/A	N/A	N/A
Maturity (years)	6.68	6.14	5.39	N/A	N/A	N/A	N/A
Term Loan vs. Credit Line	22.44%		47.02%	33.73%		47.28%	0.000
Collateral	60.03%		48.30%	49.59%		49.07%	0.000
Primary Guarantor	34.03%		47.23%	38.89%		48.75%	0.000
Primary Guarantor Monthly Salary	\$36,564	\$33,011	\$85,480	\$33,378	\$30,892	\$92,475	0.004
SBA Guarantee	0.56%		4.70%	12.21%		27.45%	0.000
Firm-Bank Dist (miles by car)	9.91	2.62	21.44	10.67	2.98	28.94	0.017
Firm-Comp Dist (miles by car)	1.10	0.55	1.59	0.93	0.48	1.48	0.000
Firm-Bank Time (minutes by car)	10.25	6.79	21.39	13.35	7.93	22.99	0.000
Firm-Comp Time (minutes by car)	2.18	1.16	4.99	2.12	1.09	4.51	1.000
Firm-Bank Aerial Distance (miles)	7.68	2.00	23.55	8.49	2.41	17.23	0.002
Firm-Comp Aerial Distance (miles)	0.74	0.38	1.58	0.68	0.34	1.15	0.001
Internal Credit Score	1036.35	1042.44	139.32	810.72	846.89	128.79	0.000
Private Credit Assessment \hat{u}_i	0.0379	0.0112	0.7224	-0.0349	-0.0106	0.5830	0.000
Scope of Banking Relationship	35.14%		44.03%	25.38%		43.52%	0.000
Months on Books	43.17	30.50	56.68	17.66	14.38	29.74	0.000
Monthly Deposit Account Balance	\$16,983	\$11,834	\$62,777	\$11,549	\$10,035	\$21,047	0.000
Months in Business	115.39	96.34	107.28	90.88	81.03	99.28	0.000
Firm's Monthly Net Income	\$110,367	\$94,724	\$256,941	\$91,350	\$84,441	\$375,803	0.000
Case-Shiller HPI	170.35	156.35	30.97	162.33	150.75	31.89	0.000
Maturity-Matched UST Yield	3.89%	3.83%	1.96%	N/A	N/A	N/A	N/A
5Y-3M UST Yield Spread (bpts)	218.92	209.24	57.65	N/A	N/A	N/A	N/A
Number of observations		12,823			12,664		25,487

This table presents summary statistics for the key variables described in Section 2 for our full sample of 25,487 data points in function of the bank's decision to offer (12,823 observations) or to deny (12,664 observations) credit to the applicant. The last column indicates the *P*-values of a two-sided *t*-test for the equality of the variables' mean conditional on the bank's decision (wherever appropriate).

zero (worst) to one thousand eight hundred fifty (best), and means (medians) in our sample are 1,036 (1,042) for successful applicants and 811 (847) for unsuccessful ones with a P -value for the t -test of $< 0.01\%$.

We capture the nature of the business ties, which might facilitate access to borrower-specific private information,¹⁰ through the length of the lending relationship (*Months on Book*: on average forty-three and eighteen months for successful and unsuccessful applicants, respectively) and its breadth. For the latter, we construct a binary variable *Scope* (35.52% of loan offers, 25.38% of rejections) in terms of the firm's current-account balance (exceeding \$5,000), prior borrowing, and the purchase of at least one other banking product (investment account, 401(k) plan, small-business credit card, mortgage services, etc.). To control for the availability of public information and firm-specific attributes, we rely on firm age (*Months in Business*: one hundred fifteen versus ninety-one months for accepted and rejected loan applications, respectively), its monthly *Net Income* (\$110,367 versus \$91,350 for successful and unsuccessful applicants, respectively), and thirty-eight industry dummy variables based on the firms' two-digit Standard-Industry Classification (SIC) codes.

Our data also contain the loan terms, such as the offered *Loan Amount* (mean and median of \$46,507 and \$39,687, respectively, in line with typical small-business lending), its *Maturity* (mean: 6.68 years), and the existence of *Collateral* (about 60% approved versus 49% rejected applications). About 34% (39%) of accepted (denied) credit applications were personally guaranteed by guarantors with a monthly income of \$36,564 (\$33,378). In case of an offer, the mean (median) of the annual percentage rate (*APR*; that is, the all-in cost of credit taking into account fees and commissions), which branches can adjust in light of local conditions and information, is 8.67% corresponding to an average (median) spread of 457 (429) basis points over maturity-matched *U.S. Treasury Yields*. Of all approved (rejected) applications, 22.44% (33.73%) are for term loans, and the remainder are for credit lines. Since our lender offers term loans only at fixed rates and credit lines at variable rates, the binary variable *Term Loan* also captures the type of interest rate.

We rely on the *Case-Shiller Home Price Index (HPI)* (Case and Shiller 1987, 1989), matching each loan application by ZIP code and month to control for potential endogeneities in the parties' choice of location and loan-transaction effects, which are due to the local level of economic activity, differences in affluence across postal zones, and differential levels of urbanization or road infrastructure as reflected in local house prices. Using IRS income and demographic U.S. Census data by ZIP code does not change the results (estimations not reported). Similarly, we rely on state and quarter dummy variables to account for regional and business-cycle effects. To control for the interest rate environment, we use the maturity-matched (interpolated)

¹⁰ James (1987) and Lummer and McConnell (1989) present evidence suggesting that banks gain access to private information over the course of the lending relationship.

U.S. Treasury Yield on the loan date and the difference between the five-year and three-month U.S. Treasury Yield (*Term Spread*).

2.3 Private information and estimation strategy

We first construct a clean measure of the bank's private information. Since each final *Proprietary Score* as a summary statistic of the bank's credit screen might also comprise publicly available data, we orthogonalize it with two widely accessible public credit-quality signals on the day of the loan request: Experian's *Commercial Intelliscore (XCI)* and the owner's personal *National Risk Model score (XNRM)*; see [Experian 2001, 2004](#) for further details). Including the latter not only mirrors established operational practices but also allows us to control for certain parts of the bank's proprietary information, which are typically correlated with personal information about manager-owners such as liens on their property pledged as security, a more precise local estimate of the value of such collateral, personal debt history, etc. To measure the purely private component of the lender's credit screen, we simply take the residual \hat{u}_i of the following regression (estimated with branch fixed effects), which we label the *Subjective Credit Assessment* ("SCA"):

$$\ln(\text{PropScore}_i) = \beta_0 + \beta_1 \cdot \text{XCI}_i + \beta_2 \cdot \text{XNRM}_i + u_i$$

Incidentally, the R^2 of the above regression is 0.73, which leaves 27% of the proprietary score's variation unexplained and, therefore, bears out the bank's contention that its own score typically comprises 20% to 30% of subjective information.¹¹ An additional benefit of the orthogonalization is the removal of hard information to the extent that the Experian scores primarily consist of objectively verifiable determinants of credit quality in the sense of [Petersen \(2004\)](#).

In terms of specifications, we simply follow the typical lending process and start with a logistic discrete-choice model of the bank's decision to offer or deny credit. We next estimate a model of the offered loan's *APR* (all-in cost). These estimations are not subject to any endogeneity problems because our sample consists of all loan applications and offers so that we can specify models of offered loan rates, not the booked ones as is common in the literature. Since firms have not yet chosen whether to accept or to decline an offer, their decision affects neither the recorded loans nor their terms. By including the 874 declined loan offers in our analysis, we therefore avoid the joint endogeneity of credit terms (such as interest rates) and shared bank-borrower attributes (such as distance) that typically arises in the study of booked loans (see, e.g.,

¹¹ For confidentiality reasons, we cannot provide further details on the orthogonalization nor report any results. Furthermore, our data provider asked us not to use the bank-internal score as a dependent variable; instead we include interaction terms in our major specifications to ascertain the effect of distance on the quality of proprietary and private information. The log-linear specification best agrees with the highly nonlinear nature of the Experian scores.

Berger, Miller, Petersen, Rajan, and Stein 2005). To address potential sample-selectivity issues, we repeat the *APR* estimations with the Heckman correction, taking into account that the credit decision precedes the structuring of the loan offer.

We explore the effect of distance on the bank's local information advantage in the face of credit-market competition by estimating a discrete-choice model of the applicant's decision to decline the offered loan terms and switch lenders. Given our hypothesis that distance affects credit decisions indirectly through its influence on the quality of credit screens, we finally specify a logistic model of borrower delinquency. Throughout the analysis, we explicitly model this complementarity between distance and private information by including appropriate interaction terms to further test our hypothesis that proximity improves the bank's intelligence. Similarly, we interact our relationship variables with the bank's private information to ascertain the source of such knowledge.¹²

We estimate all our discrete-choice specifications by full-information maximum likelihood and report their pseudo- R^2 , which is simply McFadden's likelihood ratio index, whenever appropriate. The quoted-*APR* estimations are by ordinary least squares. To account for possible systematic variations in proprietary-score revisions, credit policies, and loan pricing across branches, we rely on fixed (branch) effects and compute clustered standard errors that are adjusted for heteroscedasticity across branch offices and correlation within.¹³ As robustness checks, we also estimate all our specifications with driving minutes and aerial distances but do not tabulate the results because they are virtually identical. Since several of the variables fit better in logarithms than levels, we use the former whenever appropriate.

3. The Determinants of Credit Decisions

This section summarizes our empirical findings for credit offers and their pricing.

3.1 Loan offers

Specification 1 in Panel A of table 2 shows that the bank's credit decision depends on borrower proximity: the likelihood of obtaining credit decreases in the distance between firm and bank but increases in that between a firm and the nearest competitor. Hence, the lender is more likely to grant credit to firms in the vicinity of its own branches but less likely to do so when an applicant is close to a competitor, which is consistent with lending in the presence of location-varying adverse selection. Both effects are highly significant at the

¹² To properly assess the interaction terms' sign, marginal effect, and statistical significance in nonlinear specifications such as ours, we follow Ai and Norton (2003).

¹³ Testing for homoscedasticity, we fail to reject the hypothesis for all specifications presumably because of the low upper bound on the size of the firms in our sample (\$10 million turnover under the SME loan definition).

Table 2
The decision to offer or deny credit

Specification variable	1			2			3		
	Coeff	<i>P</i> -val	Marg	Coeff.	<i>P</i> -val	Marg	Coeff.	<i>P</i> -val	Marg
Panel A: Distance and subjective credit assessment									
Constant	2.845	0.001		3.481	0.001		1.546	0.001	
ln(1+Firm-Bank Dist)	-0.092	0.001	-4.12%	-0.208	0.0724	-0.82%	-0.669	0.2067	-0.42%
ln(1+Firm-Comp Dist)	0.265	0.001	2.20%	0.615	0.0833	0.19%	0.061	0.1969	0.36%
<i>Subjective Credit Assessment</i>									
SCA-ln(1+Firm-Bank Dist)				0.739	0.001	13.98%	0.902	0.001	13.71%
SCA-ln(1+Firm-Comp Dist)							-0.001	0.001	-1.81%
<i>Scope</i>							-0.008	0.001	-1.93%
ln(1+Months on Books)	0.710	0.001	1.03%	0.958	0.001	1.21%	0.596	0.0011	4.20%
ln(1+Months in Business)	1.567	0.001	21.57%	2.356	0.001	27.70%	1.392	0.001	10.66%
ln(1+Net Income)	0.377	0.001	13.55%	0.230	0.001	10.86%	0.003	0.001	7.32%
ln(1+Case-Shiller HPI)	0.205	0.001	3.83%	0.581	0.001	5.08%	1.929	0.0489	4.20%
<i>Collateral</i>	0.007	0.001	3.25%	0.006	0.001	2.94%	0.006	0.001	2.97%
Personal Guarantee	0.394	0.001	11.02%	1.027	0.001	12.26%	0.479	0.001	9.45%
SBA Guarantee	3.093	0.001	0.27%	2.387	0.001	0.78%	1.600	0.001	2.07%
Term Loan	-5.082	0.001	-3.23%	-4.310	0.001	-3.77%	-4.098	0.001	-3.63%
4 Quarterly Dummies	-0.947	0.029	-1.04%	-1.520	0.001	-2.43%	-3.764	0.001	-1.94%
8 State Dummies		Yes			Yes			Yes	
38 SIC Dummies		Yes			Yes			Yes	
Number of Obs		25,487			25,487			25,487	
Pseudo- <i>R</i> ²		6.74%			10.17%			10.19%	

Table 2
(Continued)

Specification Variable	1			2			3		
	Coeff.	<i>P</i> -val	Marg	Coeff.	<i>P</i> -val	Marg	Coeff.	<i>P</i> -val	Marg
Panel B: Distance and public credit-quality signals									
Constant	2.836	0.001		2.807	0.001		2.790	0.001	
ln(1+ <i>Firm-Bank Dist</i>)	-0.089	0.001	-4.02%	-0.090	0.001	-3.99%	-0.089	0.001	-4.00%
ln(1+ <i>Firm-Comp Dist</i>)	0.266	0.001	2.20%	0.267	0.001	2.20%	0.255	0.001	2.19%
ln(1+ <i>Commercial Intelliscore</i>)	0.825	0.001	9.80%				0.615	0.001	6.35%
ln(1+ <i>National Risk Model</i>)				0.458	0.001	10.48%	0.443	0.001	8.17%
ln(1+ <i>XCI</i>)-ln(1+ <i>Firm-Bank Dist</i>)							0.091	0.001	1.13%
ln(1+ <i>XCI</i>)-ln(1+ <i>Firm-Comp Dist</i>)							0.072	0.001	0.99%
ln(1+ <i>NRM</i>)-ln(1+ <i>Firm-Bank Dist</i>)							0.086	0.001	0.83%
ln(1+ <i>NRM</i>)-ln(1+ <i>Firm-Comp Dist</i>)							0.185	0.001	0.96%
<i>Scope</i>	0.717	0.001	1.03%	0.707	0.001	1.02%	0.687	0.001	1.03%
ln(1+ <i>Months on Books</i>)	1.548	0.001	21.55%	1.549	0.001	21.20%	1.559	0.001	21.55%
ln(1+ <i>Months in Business</i>)	0.373	0.001	13.69%	0.371	0.001	13.53%	0.365	0.001	13.02%
ln(1+ <i>Net Income</i>)	0.200	0.001	3.82%	0.199	0.001	3.79%	0.204	0.001	3.71%
ln(1+ <i>Case-Shiller HPI</i>)	0.007	0.001	3.25%	0.007	0.001	3.28%	0.007	0.001	3.16%
<i>Collateral</i>	0.387	0.001	10.92%	0.381	0.001	10.98%	0.383	0.001	10.77%
Personal Guarantee	3.127	0.001	0.27%	3.071	0.001	0.27%	3.028	0.001	0.27%
SBA Guarantee	-4.976	0.001	-3.17%	-4.990	0.001	-3.20%	-4.882	0.001	-3.18%
<i>Term Loan</i>	-0.932	0.029	-1.03%	-0.933	0.029	-1.03%	-0.929	0.029	-1.02%
4 Quarterly Dummies		Yes			Yes			Yes	
8 State Dummies		Yes			Yes			Yes	
38 SIC Dummies		Yes			Yes			Yes	
Number of Obs		25,487			25,487			25,487	
Pseudo- R^2		9.07%			9.06%			9.10%	

The preceding table reports the results from estimating logistic discrete-choice models of the loan application's outcome by full-information maximum likelihood with branch fixed effects and clustered standard errors that are adjusted for heteroscedasticity across branch offices and correlation within. The dependent variable is the bank's decision to offer ($Y = 1$: 12,823 observations) or to deny ($Y = 0$: 12,664 observations) credit. The explanatory variables are the firm-bank and firm-competitor distances, the lender's pure private information (Panel A: Subjective credit assessment abbreviated SCA in the interaction terms), public credit-quality signals (Panel B: *XCI* is the firm's *Commercial Intelliscore* and *NRM* is its owner's *National Risk Model score*) bank-borrower relationship characteristics, firm attributes, and various control variables (see Section 2 for a detailed description of the variables).

We tabulate the coefficients ("Coeff"), their P -values (" P -val"), and marginal effects ("Marg") for the decision to grant credit ($Y = 1$), which we obtain by simply evaluating $\frac{\partial \Pr}{\partial x_j} = \Lambda'(\mathbf{x}'_j \beta) \beta_j$ at the regressors' sample means and coefficient estimates $\hat{\beta}$. Since the probabilities of offering and denying credit sum to 1, the marginal effects for the decision to reject a loan application are simply the opposite of the reported ones. The pseudo- R^2 is McFadden's likelihood ratio index $1 - \frac{\log L}{\log L_0}$.

1% level. However, adding the subjective component of the bank's credit assessment¹⁴ (*SCA*) drastically reduces the statistical (from 1% to 10%) and economic significance of the distance variables, implying that borrower proximity and bank information might be correlated (Specification 2 in Panel A).

The apparent correlation between distance and information suggests that borrower proximity might affect credit decision indirectly through the lender's firm-specific knowledge. To explore this hypothesis, we next add interaction terms between subjective intelligence and the distance variables to the model (Specification 3, Panel A, table 2). We find that greater firm–bank distance reduces the beneficial impact of credit quality on the likelihood of obtaining a loan (negative marginal effect of the *SCA–Firm-Bank Dist* interaction term). An increase by one standard deviation in the *SCA* increases the likelihood of a credit offer by 8.72% for firms located next to the branch but only by 5.93% for those located the average 10.3-mile distance away. The greater the firm–bank distance, the less weight our lender seems to put on its private credit assessment, which is consistent with the notion that its soft-information component is primarily of a local nature. The more a branch lends outside its immediate vicinity, the harder the collection of local information becomes, rendering credit screens less precise.

In essence, borrower proximity acts as a proxy for the quality of soft private information. Consistent with this view, the distance variables become statistically insignificant in the presence of the soft-information interaction terms. To verify this conclusion, we next estimate the same specification with the firm's and its owner's public credit-quality signal, that is, Experian's *Commercial Intelliscore (XCI)* and the owner's personal *National Risk Model score (XNRM)*. Panel B of table 2 shows that, contrary to private information, the distance variables now retain their statistical and economic significance. The *SCA* as a measure of subjective intelligence does not reduce the effect of distance on credit decisions due to omitted firm characteristics. We take this counterfactual finding as strong corroboration that our *SCA* variable indeed captures soft information whose quality depends on borrower proximity. Despite advancements in borrower screening such as credit scoring, banks still rely heavily on soft information gathered through personal contacts.

Prior lending relationships, which, like borrower proximity, facilitate the collection of private information and the formation of subjective impressions of credit quality, greatly increase the likelihood of obtaining a loan (see also Agarwal, Chomsisengphet, Liu, and Souleles 2009). The uniformly high impact of their length (*Months on Books*) and, to a lesser degree, breadth (*Scope*) controlling for private (Panel A) and public (Panel B) information suggests that they represent an independent conduit for subjective intelligence to influence the outcome of credit requests. Similarly, the longer a firm has been in

¹⁴ An increase in the bank's private perception of credit quality by one standard deviation (0.64) above the *SCA*'s mean (0.0003) increases the likelihood of a credit offer by about 9%.

operation (*Months in Business*), the easier it becomes to borrow. Table 3 replicates the preceding analysis with the *Proprietary Score*, which replaces the *Subjective Credit Assessment* as a measure of the bank's overall proprietary information, confirming our previous findings.

3.2 Loan pricing

We first estimate a model of the offered loan rate (*APR*) without any private-information variables (Specification 1 in table 4). The results show that the quoted all-in cost decreases in the firm–bank distance but increases in the distance between the firm and the nearest competitor branch. A reduction of the firm–borrower distance from ten (the average) to nine miles, for instance, increases the loan rate by thirteen basis points.¹⁵ By contrast, an increase in the firm–competitor distance from one (the average) to two miles raises the quoted interest rate by fifty-six basis points. Although the bank seems to engage in spatial price discrimination tempered by the physical proximity of its customers to competitors, the results do not allow us to distinguish between direct effects of distance on offered loan rates (transportation costs) and indirect ones acting through its influence on proprietary intelligence (asymmetric information).

Including the bank's *Subjective Credit Assessment* dramatically changes the results (Specification 2). The distance variables again become statistically insignificant, confirming the previously identified association between borrower proximity and private information. Other variables measuring the bank's ability to extract information from the lending relationship (*Scope*, *Months on Books*) retain their high statistical and economic significance.¹⁶ Interacting the *SCA* with distance (Specification 3) shows that the marginal effect of a higher private signal becomes more pronounced with borrower proximity because the coefficients of the *SCA* and the *SCA–Firm-Bank Dist* interaction term are both negative. A high subjective assessment decreases the offered *APR* less for firms in the vicinity of the branch than for applicants located farther away. To put this effect into perspective, an increase in the *SCA* by one standard deviation reduces the offered *APR* by 103 basis points at the average firm–bank distance of ten miles but does so only by eighty-one basis points next to the branch. The closer firms are, the more trust the lender places in its own information and the more aggressively it can price loans because the greater threat of adverse selection deters competition.

We also find that the farther away applicants are from competitor branches, the less does an increase in perceived creditworthiness reduce the offered rate.

¹⁵ To find the marginal effects of explanatory variables in logarithms, we simply use the chain rule and, as for the logistic marginal effects, evaluate the resulting expression at the sample mean of the regressor.

¹⁶ These findings contrast with the results in Degryse and Ongena (2005) or Carling and Lundberg (2002); however, their samples are drawn from much smaller and more homogeneous populations so that informational considerations might be less important for lending decisions.

Table 3
Proprietary information in the decision to offer credit

Specification variable	1			2			3		
	Coeff	P-val	Marg	Coeff.	P-val	Marg	Coeff.	P-val	Marg
Constant	2.845	0.001		3.956	0.001		1.576	0.001	
ln(1+Firm-Bank Dist)	-0.092	0.001	-4.12%	-0.223	0.048	-1.28%	-0.742	0.149	-0.48%
ln(1+Firm-Comp Dist)	0.265	0.001	2.20%	0.625	0.105	0.20%	0.064	0.178	0.39%
ln(1+Proprietary Score)				1.327	0.001	39.97%	5.925	0.001	36.86%
ln(1+Score)-ln(1+Firm-Bank Dist)							-0.001	0.001	-1.55%
ln(1+Score)-ln(1+Firm-Comp Dist)							-0.008	0.001	-1.57%
<i>Scope</i>	0.710	0.001	1.03%	1.034	0.001	1.20%	0.640	0.001	4.69%
ln(1+Months on Books)	1.567	0.001	21.57%	2.612	0.001	27.74%	1.404	0.001	11.05%
ln(1+Months in Business)	0.377	0.001	13.55%	0.241	0.001	11.64%	0.003	0.001	7.85%
ln(1+Net Income)	0.205	0.001	3.83%	0.599	0.001	5.30%	1.932	0.049	4.70%
ln(1+Case-Shiller HPI)	0.007	0.001	3.25%	0.007	0.001	3.21%	0.007	0.001	3.26%
<i>Collateral</i>	0.394	0.001	11.02%	1.046	0.001	13.76%	0.488	0.001	10.53%
Personal Guarantee	3.093	0.001	0.27%	2.665	0.001	0.85%	1.827	0.001	2.30%
SBA Guarantee	-5.082	0.001	-3.23%	-4.735	0.001	-3.79%	-4.627	0.001	-4.13%
<i>Term Loan</i>	-0.947	0.029	-1.04%	-1.583	0.001	-2.71%	-4.036	0.001	-2.21%
4 Quarterly Dummies		Yes			Yes			Yes	
8 State Dummies		Yes			Yes			Yes	
38 SIC Dummies		Yes			Yes			Yes	
Number of Obs		25,487			25,487			25,487	
Pseudo-R ²		6.74%			9.02%			9.10%	

In the above table, we replace the bank's *Subjective Credit Assessment* with its *Proprietary Score*. See the notes to table 2 for further details and Section 2 for a description of the variables.

Table 4
Determinants of offered loan rates

Specification variable	1		2		3		4		5	
	Coeff	P-val	Coeff.	P-val	Coeff.	P-val	Coeff.	P-val	Coeff.	P-val
Constant	10.413	0.001	10.477	0.001	9.188	0.001	11.859	0.001	10.365	0.001
ln(1+Firm-Bank Dist)	-1.435	0.001	-0.980	0.190	-1.275	0.307	-1.010	0.187	-1.423	0.307
ln(1+Firm-Comp Dist)	1.173	0.001	0.597	0.241	1.321	0.405	0.619	0.238	1.310	0.489
<i>Subjective Credit Assessment</i>			-1.085	0.001	-1.127	0.001				
SCA-ln(1+Firm-Bank Dist)					-0.125	0.001				
SCA-ln(1+Firm-Comp Dist)					0.106	0.001				
ln(1+Proprietary Score)							-2.413	0.001	-2.074	0.001
ln(1+Score)-ln(1+F-B Dist)									-0.246	0.001
ln(1+Score)-ln(1+F-C Dist)									0.170	0.001
<i>Scope</i>	-1.516	0.001	-1.455	0.001	-1.312	0.001	-1.478	0.001	-1.329	0.001
ln(1+Months on Books)	-2.066	0.001	-1.871	0.001	-1.399	0.001	-2.190	0.001	-1.377	0.001
ln(1+Months in Business)	-3.079	0.001	-2.419	0.001	-1.432	0.001	-2.792	0.001	-1.441	0.001
ln(1+Net Income)	-3.514	0.001	-2.777	0.001	-1.082	0.001	-2.928	0.001	-1.105	0.001
ln(1+Case-Shiller HPI)	-0.479	0.001	-0.399	0.001	-0.433	0.001	-0.459	0.001	-0.461	0.001
<i>Collateral</i>	-2.248	0.001	-2.363	0.001	-2.337	0.001	-2.510	0.001	-2.651	0.001
Personal Guarantee	-0.880	0.001	-0.690	0.001	-1.009	0.001	-0.781	0.001	-1.159	0.001
SBA Guarantee	0.526	0.001	0.660	0.001	0.219	0.001	0.674	0.001	0.251	0.001
<i>Term Loan</i>	0.344	0.001	0.456	0.001	0.401	0.001	0.528	0.001	0.424	0.001
ln(1+Maturity)	-0.266	0.001	-0.102	0.001	-0.228	0.001	-0.119	0.001	-0.239	0.001
UST Yield	0.386	0.001	0.411	0.001	0.341	0.001	0.408	0.001	0.371	0.001
<i>Term Spread</i>	0.503	0.001	0.491	0.001	0.563	0.001	0.505	0.001	0.617	0.001
4 Quarterly Dummies	Yes		Yes		Yes		Yes		Yes	
8 State Dummies	Yes		Yes		Yes		Yes		Yes	
38 2-digit SIC Dummies	Yes		Yes		Yes		Yes		Yes	
Number of Obs	12,823		12,823		12,823		12,823		12,823	
Adjusted R ²	20.84%		24.26%		24.41%		22.79%		21.98%	

This table reports the results from regressing the offered loan rate (*APR*: all-in cost of the loan) on the firm–bank and firm–competitor driving distances in miles (abbreviated “F-B Dist” and “F-C Dist” in the interaction terms, respectively), soft and private information (*Subjective Credit Assessment* and *Proprietary Score*, respectively), bank–borrower relationship characteristics, firm attributes, and various control variables. We estimate the specifications by ordinary least squares with branch fixed effects and compute clustered standard errors that are adjusted for heteroscedasticity across branch offices and correlation within. See Section 2 for a description of the variables.

For the same eighty-one-basis-point reduction in the *APR* next to a branch office due to an increase in the *SCA* by one standard deviation, the offered rate increases by six basis points (i.e., decreases by only a net seventy-five basis points) if the applicant is located the sample mean (one mile) from the nearest competitor branch. This second interaction effect provides further evidence for the erosion of the bank's informational advantage at the periphery of local captive markets, which limits its ability to extract rents from high-quality borrowers. Replicating the analysis with the bank's *Proprietary Score* instead of its soft-information component (Specifications 4 and 5) confirms the results and shows that the information-erosion effect applies to proprietary information, too. By contrast, the distance variables remain statistically highly significant when we use the firm's and its owner's public credit scores instead (see Online Appendix). These findings further highlight how distance affects loan pricing through its effect on the quality of soft information, which, in turn, affects the strategic use of proprietary intelligence.

Having obtained credit from the bank for a year more than the average forty-three months (*Months on Books*) reduces the offered rate by up to fifty-six basis points, as does the purchase of additional products and maintenance of a deposit account with the lender (*Scope*: -151 basis points). When we interact these two relationship measures with the distance variables, the corresponding coefficients come out statistically insignificant (results not reported), which the literature has interpreted as evidence in favor of transportation-cost models of credit markets (Degryse and Ongena 2005). However, relationship variables only proxy for the ability to gather proprietary intelligence, not for the outcome of such efforts. Hence, the absence of interaction effects might also be due to the omission of variables that directly measure the bank's private information. In fact, our results point to proprietary intelligence as a conduit through which distance affects interest rate offers.

Recall that we do not need to correct for the possible joint endogeneity of loan rates and distance because our sample consists of all loan offers including the declined ones. However, this unique feature of our data also allows us to assess the degree to which an analysis of booked loans might suffer from sample-selection bias. Reestimating our loan-rate model with the Heckman correction, we achieve identification by including the income available for a personal guarantee in the selection equation. We can drop this instrument from the *APR* regression because it only affects the lender's prior credit decision but not its loan pricing. A comparison of the specifications in table 5 with the earlier results shows that this correction is important only when we do not include our proxy for private information. Once we include the *Subjective Credit Assessment*, we find that the inverse Mills ratio (Λ) becomes statistically insignificant (Specifications 2 and 3 in table 5).

The intuition is straightforward. The bank's prior decision to offer a loan is a direct function of its own credit assessment captured by the *SCA* or *Proprietary Score*. Hence, the inverse Mills ratio (Λ) and the information variables

Table 5
Determinants of offered loan rates with sample-selectivity correction

Specification variable	1		2		3	
	Coeff.	P-val	Coeff.	P-val	Coeff.	P-val
Constant	10.165	0.001	11.321	0.001	11.841	0.001
ln(1+Firm-Bank Dist)	-1.538	0.001	-1.140	0.170	-0.711	0.665
ln(1+Firm-Comp Dist)	1.431	0.001	0.752	0.234	0.323	0.751
<i>Subjective Credit Assessment</i>			-1.434	0.001	-1.295	0.001
SCA-ln(1+Firm-Bank Dist)					-0.145	0.001
SCA-ln(1+Firm-Comp Dist)					0.074	0.001
<i>Scope</i>						
ln(1+Months on Books)	-1.139	0.001	-1.217	0.001	-1.212	0.001
ln(1+Months in Business)	-4.244	0.001	-4.325	0.001	-2.160	0.001
ln(1+Net Income)	-1.675	0.001	-1.578	0.001	-2.454	0.001
ln(1+Case-Shiller HPI)	-2.621	0.001	-2.597	0.001	-2.539	0.001
<i>Collateral</i>						
Personal Guarantee	-0.344	0.001	-0.377	0.001	-0.379	0.001
SBA Guarantee	-1.875	0.001	-1.821	0.001	-2.743	0.001
Term Loan	-0.487	0.001	-0.463	0.001	-0.481	0.001
ln(1+Maturity)	0.206	0.001	0.199	0.024	0.797	0.001
UST Yield	0.943	0.001	0.934	0.019	1.058	0.001
Term Spread	-0.042	0.001	-0.042	0.001	-0.054	0.001
Lambda	0.455	0.001	0.517	0.001	0.736	0.001
4 Quarterly Dummies	0.612	0.001	0.646	0.001	0.492	0.001
8 State Dummies	4.589	0.001	3.734	0.125	3.363	0.170
38 2-digit SIC Dummies	Yes		Yes		Yes	
Number of Obs	Yes		Yes		Yes	
Adjusted R ²	12,823		12,823		12,823	
	20.78%		23.77%		24.01%	

This table reports the results from regressing the offered loan rate (*APR*: all-in cost of the loan) on the firm–bank and firm–competitor driving distances in miles, subjective intelligence, bank–borrower relationship characteristics, firm attributes, and various control variables correcting for the bank’s prior credit decision, which might result in sample-selection bias. To this end, we include Lambda, the inverse Mills ratio (hazard rate) of the Probit selection model, required by the Heckman correction. The selection equation includes the monthly income available for a personal guarantee as the identifying instrument, which we drop in the main estimation. Again, we estimate the specifications with branch fixed effects and use clustered standard errors that are adjusted for heteroscedasticity across branch offices and correlation within. Section 2 contains a detailed description of the variables.

should be correlated, explaining the loss of statistical significance of the former when we add the latter. Our results suggest two different avenues to address sample-selectivity problems typically arising in the analysis of booked loans. One can correct for selection bias either indirectly using the Heckman procedure or directly with the information variables underlying the prior (credit) decision that Lambda controls for. In the absence of information variables, the correction factor Lambda is statistically significant, and its inclusion increases the effect of distance between applicant and competitor branch on loan pricing (Specification 1, table 5). We interpret these findings as evidence for a mild sample-selection bias arising from the bank’s prior credit decision for offered and, *a fortiori*, booked loans. A comparison of tables 4 and 5 shows that all other effects remain more or less unchanged.

3.3 Distance and the availability and pricing of credit

Our findings point to the following fundamental trade-off: borrower proximity facilitates access to credit through branch offices' local knowledge of firms and economic conditions (Specification 3, tables 2 and 3), albeit at the price of higher interest rates (Specification 3, table 4). Consequently, distance affects loan offers and terms indirectly through the ease of access to local information, which improves the quality of credit screens through the incorporation of subjective credit assessments.¹⁷ The results reveal that such private information is particularly valuable because it allows banks to soften competition and to extract location-varying rents in informationally sensitive markets such as small-business lending. Subjective intelligence is at the root of a local information advantage, which creates an adverse-selection threat and therefore reduces the effect of a competitor's geographic reach on lending decisions and loan terms.

When we carry out the same analysis with public measures of credit quality, we see that our distance variables do not lose their statistical and economic significance although we control for the length and scope of the lending relationship, which are typically associated with the production of private information. This counterfactual exercise shows that proprietary intelligence and, especially, its subjective component more likely than not depend on borrower proximity, which points to the local nature of soft information. To assess the robustness of this interpretation, we also regress our public credit signals (the firm's Business Intelliscore and its owner's *National Risk Model score*) and the lender's proprietary score on a constant and the two distance variables (see Online Appendix), which are statistically significant for the latter but not the former. When we estimate a linear model of the SCA's standard deviation as a function of the average firm–bank and firm–competitor distances by branch, we find that it decreases in borrower proximity, providing direct evidence that the quality of soft information varies with borrower location (see Online Appendix).

Two conclusions follow. Firstly, the lender's private information seems intimately related to its knowledge of the local economy. Since both the *Proprietary Score* and, *a fortiori*, the *Subjective Credit Assessment* contain personal impressions of credit quality through score revisions by loan officers, we would expect the bank's private information to primarily consist of soft local intelligence. Secondly, such private information seems to weaken the effect of a competitor's proximity on the decision to grant credit and its pricing. This finding is consistent with the notion that soft information can limit competition by counteracting either a transportation cost or any informational advantage, which a competitor might hold in turn, with an adverse-selection threat.

¹⁷ Agarwal and Hauswald (2007) provide evidence that firms disclose confidential information, which facilitates the acquisition of soft intelligence, when seeking to establish long-term lending relationships.

4. Spatial Competition and Choice of Lender

When the severity of information asymmetries decreases with physical distance, a key result holds that transactions become more contested as the informational advantage of the better informed party declines (Hauswald and Marquez 2006). If firm proximity was to improve the quality of the bank's information, we would expect the competition for borrowers to increase with distance so that borrowers located farther away should switch lenders more frequently. By contrast, borrowers receive at most one credit offer in transportation-cost models and, therefore, do not have any incentive to switch banks, which is clearly at odds with our data.

To investigate the competitive reaction in local credit markets, we estimate a logistic discrete-choice model of the borrower's decision to decline a loan. By comparing offers with actually booked loans, we identify 874 instances (out of 12,823: approximately 6.8%) in which firms refused the bank's terms and sought credit from a competitor around the same time.¹⁸ Matching these observations with credit-bureau information indicates that 539 firms (61.7%) indeed accepted the competing loan offer. The remainder (335 firms: 38.3%) declined the competing offer to approach yet another lender, failed to agree on terms, or withdrew from the market.

Specification 1 in table 6 shows that more distant firms are more likely to decline the bank's loan offer. By contrast, potential borrowers located farther away from the nearest competitor are less likely to refuse a loan offer, and both distance variables are statistically highly significant. Moreover, the longer the lending relationship and, presumably, the better the firm is known to the bank (*Months on Books*, *Scope*), the less likely it is to refuse a loan offer and switch lenders. Borrowing costs also affect the applicant's decision: the higher the quoted *APR*, the more likely the firm is to decline the offer, presumably because it hopes to get better terms elsewhere. These results are broadly consistent with competition for borrowers, which increases with distance. On the one hand, longer lending relationships or greater proximity to the bank increases the threat of adverse selection for competitors, thereby reducing competition. On the other hand, attempts to extract rents through higher loan rates, especially over greater distances, facilitate the poaching of good credit risks as reflected in more frequently declined credit offers.

Once we include the bank's *Subjective Credit Assessment* (Specification 2, table 6), distance becomes statistically insignificant. At the same time, our proxy for the bank's private information is highly significant. The higher the *SCA* and, hence, the perceived credit quality, the more likely an applicant is to decline the loan offer. Not only is it easier for better credit risks to obtain competing loan offers but also they are the primary targets for rent extraction (see, e.g., von Thadden 2004) and, hence, have a larger incentive to switch lenders.

¹⁸ The decision to switch lenders is very different from a borrower's choice of single versus multiple banking relationships; see Detragiache, Garella, and Guiso (2000) and Farinha and Santos (2002).

Table 6
The decision to decline loan offers

Specification variable	Coeff	1			2			3		
		P-val	Marg	Coeff.	P-val	Marg	Coeff.	P-val	Marg	
Constant	-2.075	0.001		-2.496	0.001		-2.385	0.001		
ln(1+Firm-Bank Dist)	2.587	0.001	5.50%	1.322	0.278	0.61%	0.408	0.604	1.11%	
ln(1+Firm-Comp Dist)	-1.038	0.001	-7.31%	-1.280	0.384	-0.86%	-0.556	0.590	-0.05%	
<i>Subjective Credit Assessment</i>				2.036	0.001	8.02%	1.810	0.001	10.38%	
SCA-ln(1+Firm-Bank Dist)							0.755	0.001	1.21%	
SCA-ln(1+Firm-Comp Dist)							-0.128	0.237	-0.92%	
<i>Scope</i>	-0.971	0.001	-8.33%	-1.301	0.001	-11.81%	-1.633	0.001	-9.32%	
ln(1+Months on Books)	-1.002	0.001	-5.90%	-1.249	0.001	-5.98%	-0.934	0.001	-5.31%	
ln(1+Months in Business)	-4.110	0.001	-3.30%	-1.381	0.001	-4.44%	-1.017	0.001	-12.03%	
ln(1+Net Income)	2.977	0.001	6.85%	3.688	0.001	6.56%	0.949	0.001	3.88%	
ln(1+Case-Shiller HPI)	0.002	0.476	0.24%	0.002	0.727	0.23%	0.002	0.730	0.24%	
<i>Collateral</i>	0.165	0.001	2.77%	0.314	0.545	4.00%	0.236	0.544	1.11%	
Personal Guarantee	4.967	0.001	5.47%	5.295	0.001	6.56%	3.535	0.001	4.73%	
SBA Guarantee	0.208	0.574	0.56%	0.429	0.470	0.44%	0.128	0.940	0.46%	
<i>Term Loan</i>	-0.189	0.504	-0.17%	-0.054	0.990	-0.25%	-1.291	0.528	-0.02%	
APR	0.318	0.001	11.90%	0.271	0.001	9.59%	0.708	0.001	9.86%	
ln(Loan Amount)	-10.119	0.001	-4.14%	-7.028	0.001	-5.15%	-3.425	0.001	-3.38%	
ln(1+Maturity)	-0.106	0.001	-2.49%	-0.120	0.001	-2.17%	-0.443	0.001	-0.88%	
UST Yield	1.131	0.001	10.27%	2.218	0.001	10.40%	0.366	0.001	4.36%	
<i>Term Spread</i>	-1.436	0.001	-3.44%	-1.722	0.001	-0.82%	-1.664	0.001	-1.31%	
4 Quarterly Dummies		Yes			Yes			Yes		
8 State Dummies		Yes			Yes			Yes		
38 SIC Dummies		Yes			Yes			Yes		
Number of Obs		12,823			12,823			12,823		
Pseudo- R^2		3.27%			3.42%			3.43%		

This table reports the results from estimating a logistic discrete-choice model of the borrower's decision to refuse the bank's loan offer by full-information maximum likelihood for the subsample of successful loan applications (12,823 observations) with branch fixed effects and clustered standard errors that are adjusted for heteroscedasticity across branch offices and correlation within. The dependent variable is the applicant's decision to decline ($Y = 1$: 874 observations) or to accept ($Y = 0$: 11,949 observations) the bank's offer. The explanatory variables are the firm-bank and firm-competitor driving distances in miles, private information, bank-borrower relationship characteristics, firm attributes, and various control variables. See Section 2 for a description of the variables and the notes to table 2 for further details.

When we add the *SCA*–distance interaction terms (Specification 3, table 6), we see that distance in the presence of competition weakens the bank’s informational hold over borrowers. The farther away a successful applicant is from its branch, the more the likelihood of declining a loan offer increases in its credit score. Hence, distance exacerbates the tendency of good credit risks to seek credit elsewhere, thereby indicating that other lenders indeed compete more aggressively for such borrowers. Using the *Proprietary Score* as a mixture of private and public information only strengthens the results (not reported) because firms more likely than not are aware of their own credit quality and act accordingly.

Our results are again consistent with strategic lending by intermediaries, which use soft private information to deter competition.¹⁹ The better the bank’s intelligence, that is, the better the quality of its credit screen and the greater its ability to gather soft information due to borrower proximity, the easier it becomes to extract rents because our lender has a larger informational advantage over its competitors.²⁰ Such attempts, however, fail at the branch’s periphery in terms of geographic reach or applicant quality explaining why greater *Firm-Bank Dist* or a higher credit score increases the likelihood that the applicant declines the credit offer. Taken together, we interpret these results as evidence in favor of lending strategies, which carve out and protect local captive markets through the gathering of soft private information. We find next to no support for the predictions of transportation-cost models of financial intermediation.

5. Distance and Credit Delinquency

The results so far suggest that banks use private information to create an adverse-selection threat, which allows them to limit competition for high-quality borrowers. The crucial determinants of the lender’s informational advantage are borrower proximity and the intensity of the lending relationship, which improve the precision of its credit assessment through subjective intelligence. To the extent that the relevant information is primarily local (Petersen and Rajan 2002), we would expect firm–bank distance to determine a lender’s ability to successfully assess credit risks. Hence, we can directly test our hypothesis that distance affects the quality of the bank’s private information by studying the likelihood of lending to a bad credit risk (type II error), which should increase in firm–bank distance if the precision of credit screens were to decrease in it.

Given that banks presumably price loans to compensate for differential credit quality, we test this hypothesis under the assumption that our lender

¹⁹ See Sharpe (1990), Rajan (1992), or von Thadden (2004) on this point.

²⁰ For evidence on the winner’s curse in banking, see Shaffer (1998). Schenone (2009) similarly finds limits to rent extraction: lenders exploit their informational advantage for privately owned borrowers but IPOs and the concurrent information release weaken banks’ hold over their clients.

sets interest rates optimally. Our sample contains 319 loans out of the 11,949 booked ones (approximately 2.7%), which became sixty days past due within eighteen months of the loan's origination.²¹ We choose this window so that the likelihood of a loan becoming overdue is related more to the initial credit assessment than to subsequent economic events beyond the bank's control. For comparability, we also trace the ultimate loan performance for the 12,664 unsuccessful loan applicants from credit-bureau information. About half of those were able to obtain credit from other sources in the month after the loan request's rejection. Roughly 25% of these loans became delinquent in the eighteen-month period following origination so that the delinquency rate is about ten times higher for rejected applications than for booked loans. By contrast, about 3% of successful loan applicants that switched lenders defaulted on their loans from other institutions, which is very comparable to our booked loans. Our data provider clearly minimizes type II error (lending to a bad credit risk), albeit at the expense of rejecting acceptable credit risks (type I error), but obviously cannot eliminate the former in equilibrium.

Table 7 reports the results of a logistic model of credit delinquency as a function of our usual variables. We see that both the firm–bank and firm–competitor distance enter significantly into explaining the occurrence of nonperforming loans (Specification 1). The farther a borrower is located away from the bank, the more likely eventual default becomes. In line with the empirical predictions in Hauswald and Marquez (2006), type II errors in lending increase with distance presumably because of increasingly difficult access to soft local information. This result is also consistent with the evidence in DeYoung, Glennon, and Nigro (2006), who find in their analysis of loans guaranteed by the Small Business Administration (SBA) that default increases in the distance between bank and borrower ZIP codes.

However, both distance variables become statistically insignificant when we include the bank's *Subjective Credit Assessment* (Specification 2 in table 7). Once again, borrower proximity only matters insofar as it captures the quality of the bank's credit screen and, hence, expected future loan performance, which we interpret as further evidence in favor of the local-information hypothesis. Although the relationship variables retain their significance and marginal effects, the bank's private information has by far the largest impact on predicting delinquency: the lower the firm's perceived credit quality, the more likely default becomes. Interacting the *SCA* and distance variables (Specification 3 in table 7) provides further evidence that distance affects the quality of bank information. We see that the farther away the borrower, the smaller the marginal impact of lower credit quality becomes in terms of default probabilities. A reduction in the *SCA* by one standard deviation

²¹ Although the technical definition of default is 180 days past due, most lenders, including ours, take action after at most sixty days past due, either writing off the loan, selling it off, or assigning it for collection. Over 90% of loans sixty days overdue eventually suffer default according to our data provider.

Table 7
The likelihood of credit delinquency

Specification variable	1			2			3		
	Coeff	P-val	Marg	Coeff.	P-val	Marg	Coeff.	P-val	Marg
Constant	-4.963	0.001		-5.659	0.001		-5.612	0.001	
ln(1+Firm-Bank Dist)	0.148	0.001	1.03%	0.273	0.127	0.36%	0.568	0.376	0.32%
ln(1+Firm-Comp Dist)	-0.091	0.001	-1.89%	-0.448	0.382	-0.15%	-0.499	0.207	-0.68%
<i>Subjective Credit Assessment</i>				-1.261	0.001	-14.35%	-1.198	0.001	-12.32%
SCA-ln(1+Firm-Bank Dist)							0.063	0.001	5.20%
SCA-ln(1+Firm-Comp Dist)							-0.030	0.001	-0.05%
<i>Scope</i>									
ln(1+Months on Books)	-0.819	0.001	-8.30%	-0.543	0.001	-10.46%	-0.579	0.001	-4.45%
ln(1+Months in Business)	-3.025	0.001	-7.52%	-3.618	0.001	-8.88%	-3.395	0.001	-3.68%
ln(1+Net Income)	-0.453	0.001	-4.13%	-0.482	0.001	-5.36%	-0.562	0.001	-5.77%
ln(1+Case-Shiller HPI)	-1.796	0.001	-1.68%	-2.663	0.001	-1.54%	-1.530	0.001	-8.37%
<i>Collateral</i>									
Personal Guarantee	-0.088	0.001	-5.87%	-0.081	0.001	-5.78%	-0.084	0.001	-5.81%
SBA Guarantee	-2.058	0.001	-3.33%	-3.289	0.001	-1.75%	-1.782	0.001	-3.72%
Term Loan	-3.130	0.001	-6.18%	-5.154	0.001	-6.63%	-0.937	0.001	-4.59%
APR	6.644	0.001	4.65%	3.991	0.001	4.97%	4.914	0.001	5.00%
ln(Loan Amount)	0.104	0.001	3.88%	0.478	0.001	3.83%	0.393	0.001	3.86%
ln(1+Maturity)	1.862	0.001	5.77%	2.631	0.001	7.21%	2.216	0.001	7.89%
UST Yield	-3.102	0.001	-10.65%	-2.212	0.001	-11.42%	-2.395	0.001	-14.50%
Term Spread	-0.201	0.623	-0.93%	-0.154	0.704	-5.03%	-0.337	0.773	-5.31%
4 Quarterly Dummies	0.344	0.001	2.34%	0.869	0.001	0.62%	0.206	0.001	0.47%
8 State Dummies	1.513	0.001	3.82%	1.639	0.001	4.17%	1.669	0.001	1.28%
38 SIC Dummies		Yes			Yes			Yes	
Number of Obs		Yes			Yes			Yes	
Pseudo-R ²		Yes			Yes			Yes	
		11,949			11,949			11,949	
		13.74%			17.05%			17.93%	

This table reports the results from estimating a logistic model of the likelihood that a loan becomes sixty days overdue within eighteen months of origination by full-information maximum likelihood for the subsample of actual loans booked by the bank (11,949 observations). The dependent variable is the performance status of the loan during its first eighteen months: at most sixty days overdue (corresponding to our bank's internal definition of a delinquent loan $Y = 1$: 319 observations), or current ($Y = 0$: 11,630 observations). The explanatory variables are the firm-bank and firm-competitor driving distances in miles, private information, bank-borrower relationship characteristics, firm attributes, and various control variables. See Section 2 for a description of the variables and the notes to table 2 for further details.

increases the likelihood of default by 7.8% next to the branch, as opposed to reducing it by 0.2% at the mean bank–borrower distance of ten miles.

When branches lend to borrowers outside their immediate proximity, they are more likely to fall victim to credit delinquency. Type II errors in credit decisions, which are ultimately unavoidable in equilibrium, increase with distance so that banks justifiably discount their private information for far away borrowers. These findings confirm that a lender’s informational advantage seems to result from local knowledge which the bank–borrower distance erodes. At the same time, the length and breadth of the lending relationship (*Months on Books*, *Scope*), which are both proxies for the bank’s ability to gather soft proprietary intelligence, strongly decrease the likelihood of credit delinquency. These results reinforce our conclusion that a bank’s proprietary intelligence has an important local component whose quality increases with borrower proximity. As a consequence, distance becomes a proxy for the precision of credit screens.

6. Discussion

We study the origins of private information and its strategic use in informationally sensitive markets. Taking advantage of recent innovations in credit assessments, we set our investigation in the market for loans to small firms. This setting has the dual advantage of offering just the right degree of informational opacity for private-information production to be a genuine competitive advantage and yet to be measurable through the lender’s proprietary credit scoring methodology. Having extracted the purely private component of such credit assessments, we show that both borrower proximity and pre-existing business ties act as complements in information production because they facilitate the collection and interpretation of local subjective intelligence. Geographic proximity improves the quality of private information, whereas closeness in the relationship sense provides access to the crucial soft intelligence captured in the subjective revisions of a firm’s credit score.

Our results reveal that firm–bank distance is an excellent proxy for a lender’s informational advantage, which it uses to create adverse-selection problems for its nearby competitors, to carve out local captive markets, and to partially fend off competition for its core market (see, e.g., [Gehrig 1998](#); [Dell’Ariccia 2001](#)). However, distance erodes the quality of soft local information so that competitors contest captive markets more vigorously by poaching customers at the bank’s periphery and limiting its ability to earn information rents. Technological progress can only partially substitute for the natural limits to local-information gathering over greater distances. Hence, the hardening of soft information ([Petersen 2004](#)) does not fundamentally alter the competitive dynamics in local credit markets but rather changes our perception of “local” lending.

We also identify the source of a bank's informational advantage, which crucially depends on its ability to gather soft location-specific information. In the past, banks typically would not quantify subjective impressions of credit quality but relied on loan officers' personal knowledge of the applicants to grant "character loans." However, technological progress, coupled with appropriate operational policies, means that even large financial institutions can compete in informationally difficult markets. By hardening locally gathered soft information, large lenders can bring their scale and superior data processing to bear in successfully challenging the hold of smaller players on such markets (see, e.g., [Berger, Miller, Petersen, Rajan, and Stein 2005](#)). In this sense, we provide the first evidence on how a lender's ability to harden soft information is shifting the industry's competitive dynamics and affecting its structure.

Our findings also suggest that, despite important technological advances, distance still affects information acquisition through its influence on soft intelligence in informationally sensitive markets. Local underwriters of municipal debt can price informationally opaque issues more aggressively precisely because their knowledge of local conditions and personal relationships with issuers and investors grants them access to critical soft information ([Butler 2008](#)). [Goetzman, Pons-Sanz, and Ravid \(2004\)](#), who study the role of soft information in the pricing of intellectual property, argue that, although quantification and measurement are important business tools, cognitive processing and human judgment are indispensable. However, our work shows that it is precisely the quantification of subjective impressions and their incorporation into statistical analysis that allows organizations to extend their reach and contest markets at their periphery more vigorously.

Soft information and distance also matter for the design of organizations and corporate decision making. Due to its inherently ambiguous nature, soft information is less easily transferable so that geographic dispersion can impede its flow ([Landier, Nair, and Wulf 2009](#)). In this situation, local units should have more autonomy and authority in their decision making to provide appropriate incentives to local managers for the collection, analysis, and use of subjective intelligence ([Aghion and Tirole 1997](#)). Once again, credit markets offer the perfect setting for an investigation of this hypothesis (see also [Stein 2002](#)) because loan officers accumulate soft local knowledge through personal contacts, quantify it through subjective revisions to credit scores, and communicate the hardened soft information to headquarters in case of local overrides of bank-wide lending standards. We leave this question for future research.

Appendix

Panels A and B in table 2 of the article's published version contrast the very different influence of distance on credit decisions in the presence of soft private and hard public information. In particular, we find that the distance effects preserve their full statistical and economic significance under the inclusion of public credit-quality signals. Both the firm's *Commercial Intelliscore* and its owner's *National Risk Model score* do not appear to be correlated with borrower proximity

Table A1
Public information and offered loan rates

Specification variable	1		2		3		4		5	
	Coeff.	P-val	Coeff.	P-val	Coeff.	P-val	Coeff.	P-val	Coeff.	P-val
Constant	10.053	0.001	10.132	0.001	10.277	0.001	10.402	0.001	10.203	0.001
$\ln(1+\text{Firm-Bank Dist})$	-1.417	0.001	-1.427	0.001	-1.403	0.001	-1.385	0.001	-1.385	0.001
$\ln(1+\text{Firm-Comp Dist})$	1.144	0.001	1.143	0.001	1.159	0.001	1.179	0.001	1.158	0.001
$\ln(1+\text{Commercial Intelliscore})$	-0.644	0.001			-0.503	0.001			-0.494	0.001
$\ln(1+\text{National Risk Model})$			-0.729	0.001			-0.569	0.001	-0.562	0.001
$\ln(1+XCI)-\ln(1+\text{Firm-Bank Dist})$					-0.097	0.001			-0.095	0.001
$\ln(1+XCI)-\ln(1+\text{Firm-Comp Dist})$					-0.069	0.073			-0.069	0.072
$\ln(1+NRM)-\ln(1+\text{Firm-Bank Dist})$							-0.056	0.039	-0.055	0.039
$\ln(1+NRM)-\ln(1+\text{Firm-Comp Dist})$							-0.065	0.001	-0.065	0.001
<i>Scope</i>	-1.471	0.001	-1.475	0.001	-1.505	0.001	-1.515	0.001	-1.490	0.001
$\ln(1+\text{Months on Books})$	-2.043	0.001	-2.057	0.001	-2.020	0.001	-2.038	0.001	-2.003	0.001
$\ln(1+\text{Months in Business})$	-3.037	0.001	-3.061	0.001	-3.001	0.001	-3.028	0.001	-2.973	0.001
$\ln(1+\text{Net Income})$	-3.465	0.001	-3.480	0.001	-3.552	0.001	-3.550	0.001	-3.502	0.001
$\ln(1+\text{Case-Shiller HPI})$	-0.474	0.001	-0.478	0.001	-0.479	0.001	-0.481	0.001	-0.476	0.001
<i>Collateral</i>	-2.287	0.001	-2.244	0.001	-2.194	0.001	-2.215	0.001	-2.192	0.001
Primary Guarantor	-0.860	0.001	-0.857	0.001	-0.876	0.001	-0.882	0.001	-0.875	0.001
SBA Guarantee	0.515	0.001	0.517	0.001	0.529	0.001	0.528	0.001	0.519	0.001
Term Loan	0.342	0.001	0.342	0.001	0.345	0.001	0.343	0.001	0.342	0.001
$\ln(1+\text{Maturity})$	-0.259	0.001	-0.256	0.001	-0.264	0.001	-0.260	0.001	-0.260	0.001
UST Yield	0.381	0.001	0.386	0.001	0.380	0.001	0.376	0.001	0.375	0.001
Term Spread	0.496	0.001	0.498	0.001	0.488	0.001	0.494	0.001	0.485	0.001
4 Quarterly Dummies		Yes		Yes		Yes		Yes		Yes
8 State Dummies		Yes		Yes		Yes		Yes		Yes
38 2-digit SIC Dummies		Yes		Yes		Yes		Yes		Yes
Number of Obs		12,823		12,823		12,823		12,823		12,823
Adjusted R^2		22.59%		22.57%		22.70%		22.71%		22.73%

This table reports the results from regressing the offered loan rate (*APR*: all-in cost of the loan) on the firm's *Commercial Intelliscore* and its owner's *National Risk Model score* instead of private and proprietary information (*Subjective Credit Assessment* and *Proprietary Score*, respectively). For further details, see the notes to table 5 in the article.

(Panel B). In contrast, Panel A suggests that distance and private information are correlated. Taken together, these results allow us to argue that distance affects credit decisions indirectly through the collection of subjective intelligence, which borrower proximity facilitates.

This Online Appendix further explores the differential effects of distance in credit decision when controlling for private and public information. Table A1 below replicates the public-information analysis for our loan-pricing estimations (table 4 in the published version) by replacing the *Subjective Credit Assessment (SCA)* with the public credit-quality signals, that is, the firm's *Commercial Intelliscore* and its owner's *National Risk Model score*. Comparing table 4 in the article and the attached table A1 shows again that, contrary to the inclusion of the *SCA*, the geographic effects retain their statistical and economic significance in the presence of public information. In complete analogy to the credit-decision model, the distance variables exert an independent effect on the pricing of loans and, therefore, appear to be uncorrelated with public credit-quality signals. Once again, it seems reasonable to conclude that firm proximity influences the pricing of credit through its influence on soft information as measured by the *SCA*.

Taken together, these additional results strongly suggest that the *Proprietary Score* and, hence, the *Subjective Credit Assessment* are not proxying for hard information or omitted firm characteristics, which the control variables fail to pick up. Instead, they seem to contain genuine soft information, whose quality varies in bank–borrower distance. This counterfactual exercise lends additional credence to our interpretation that firm proximity affects subjective intelligence, which, in turn, influences credit decisions.

To directly test this hypothesis, we also estimate linear models of the firm's *Commercial Intelliscore*, its owner's *National Risk Model score*, and the bank's *Proprietary Score* as a function of a constant as well as the *Firm-Bank* and *Firm-Competitor* distances. For legal and confidentiality reasons, we cannot report detailed results but, instead, summarize our findings. The two public credit-quality variables do not appear to be correlated with borrower proximity (P -values of 0.992 to 0.999 for the coefficients of the distance variables, R^2 of 1.30% and 1.20% for the *Commercial Intelliscore* and *National Risk Model* regressions, respectively). In contrast, distance affects the bank's proprietary credit assessment in a statistical and economically significant manner (P -value of 0.001 for the coefficients of the distance variables, R^2 of 15.00%).

Given the very different results for the firm's *Commercial Intelliscore* or its owner's *National Risk Model score*, as opposed to the bank's *Proprietary Score*, we would expect the latter's geographic dependence to arise from the subjective score revisions. To directly test this conjecture and our hypothesis that distance erodes the quality of private information, we finally compute the standard deviation of each branch's *Subjective Credit Assessment* over the sample period and regress it on the average *Firm-Bank* and *Firm-Competitor* driving distances by branch. We find that, consistent with our conjecture, the dispersion of the PCA indeed increases in the bank–borrower distance (P -value of 0.001, R^2 of 20.10%). These additional results provide further support for our contention that soft information depends in part on physical proximity.

References

- Agarwal, S., S. Chomsisengphet, C. Liu, and N. Souleles. 2009. Relationship Lending: Evidence from the Consumer Credit Market. Mimeo, Wharton.
- Agarwal, S., and R. Hauswald. 2007. The Choice between Arm's-Length and Inside Debt. Mimeo, FRB of Chicago and American University.
- Aghion, P., and J. Tirole. 1997. Formal and Real Authority in Organizations. *Journal of Political Economy* 105:1–29.
- Ai, C., and C. Norton. 2003. Interaction Terms in Logit and Probit. *Economic Letters* 80:123–29.
- Berger, A., W. Frame, and N. Miller. 2005. Credit Scoring and the Availability, Price, and Risk of Small Business Credit. *Journal of Money, Credit and Banking* 37:191–222.

- Berger, A., N. Miller, M. Petersen, R. Rajan, and J. Stein. 2005. Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks. *Journal of Financial Economics* 76:237–69.
- Broecker, T., 1990. Credit-Worthiness Tests and Interbank Competition. *Econometrica* 58:429–52.
- Butler, A. 2008. Distance Still Matters: Evidence from Municipal Bond Underwriting. *Review of Financial Studies* 21:763–84.
- Carling, K., and S. Lundberg. 2002. Bank Lending, Geographical Distance, and Credit Risk: An Empirical Assessment of the Church Tower Principle. Sveriges Riksbank Working Paper 144.
- Case, K. E., and R. J. Shiller. 1987. Prices of Single-Family Homes Since 1970: New Indexes for Four Cities. *New England Economic Review* (September/October):45–56.
- . 1989. The Efficiency of the Market for Single-Family Homes. *American Economic Review* 79:125–37.
- Chiappori, P.-A., D. Perez-Castrillo, and T. Verdier. 1995. Spatial Competition in the Banking System: Localization, Cross-subsidies, and the Regulation of Deposit Rates. *European Economic Review* 39:889–918.
- Coval, J., and T. Moskowitz. 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. *Journal of Finance* 54:2045–73.
- . 2001. The Geography of Investment: Informed Trading and Asset Prices. *Journal of Political Economy* 109:811–41.
- Degryse, H., and S. Ongena. 2005. Distance, Lending Relationships, and Competition. *Journal of Finance* 60:231–66.
- Dell’Ariccia, G. 2001. Asymmetric Information and the Structure of the Banking Industry. *European Economic Review* 45:1957–80.
- Detragiache, E., P. Garella, and L. Guiso. 2000. Multiple versus Single Banking Relationships: Theory and Evidence. *Journal of Finance* 55:1133–61.
- DeYoung, R., D. Glennon, and P. Nigro. 2006. Borrower–Lender Distance, Credit Scoring, and the Performance of Small Business Loans. FDIC Center for Financial Research Working Paper No. 2006-04.
- Ely, D., and K. Robinson. 2001. Consolidation, Technology, and the Changing Structure of Banks’ Small Business Lending. *Economic and Financial Review*, First Quarter 2001, Federal Reserve Bank of Dallas.
- Experian. 2001. *Commercial Intelliscore*. Experian Information Solutions, Inc., October 2001. <http://www.experian.com>.
- . 2004. *Business Owner Profile*. Experian Information Solutions, Inc., May 2004. <http://www.experian.com>.
- Farinha, M., and J. Santos. 2002. Switching from Single to Multiple Bank Lending Relationships: Determinants and Implications. *Journal of Financial Intermediation* 11:124–51.
- Freixas, X., and J.-C. Rochet. 2008. *Microeconomics of Banking*, 2nd ed. Cambridge, MA: MIT Press.
- Garmaise, M. J., and T. J. Moskowitz. 2004. Confronting Information Asymmetries: Evidence from Real-Estate Markets. *Review of Financial Studies* 17:405–37.
- Goetzman, W., V. Pons-Sanz, and S. A. Ravid. 2004. Soft Information, Hard Sell: The Role of Soft Information in the Pricing of Intellectual Property—Evidence from Screenplays Sales. Mimeo, Yale University.
- Gehrig, T. 1998. Screening, Cross-border Banking, and the Allocation of Credit. *Research in Economics* 52:387–407.
- Hau, H. 2001. Location Matters: An Examination of Trading Profits. *Journal of Finance* 56:1959–83.
- Hauswald, R., and R. Marquez. 2006. Competition and Strategic Information Acquisition in Credit Markets. *Review of Financial Studies* 19:967–1000.

- James, C. 1987. Some Evidence on the Uniqueness of Bank Loans. *Journal of Financial Economics* 19:217–35.
- Landier, A., V. Nair, and J. Wulf. 2009. Tradeoffs in Staying Close: Corporate Decision-Making and Geographic Dispersion. *Review of Financial Studies* 22:1119–48.
- Lederer, P. J., and A. P. Hurter. 1986. Competition of Firms: Discriminatory Pricing and Location. *Econometrica* 54:623–40.
- Lummer, S., and J. McConnell. 1989. Further Evidence on the Bank Lending Process and the Capital Market Response to Bank Loan Agreements. *Journal of Financial Economics* 25:99–122.
- Mara, J. 2004. SmartView Launches on Yahoo! Maps. *ClickZ Internet Advertising News* March 9, 2004. <http://www.clickz.com/news/print.php/3323441> (accessed May 2007).
- Petersen, M. 2004. Information: Hard and Soft. Mimeo, Northwestern University.
- Petersen, M., and R. Rajan. 2002. Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance* 57:2533–70.
- Ragozzino, R., and J. Reuer. 2009. Geographic Distance and M&A Markets: IPOs as Information Diffusion Mechanisms. *Strategic Management Journal*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=785915.
- Rajan, R. 1992. Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt. *Journal of Finance* 47:1367–400.
- SBA. 2005. *Small Business and Micro Business Lending in the United States for Data Years 2002–2003*. Small Business Administration Office of Advocacy, March 2005.
- Schenone, C. 2009. Lending Relationships and Information Rents: Do Banks Exploit Their Information Advantages? *Review of Financial Studies*. Advance Access published October 20, 2009, 10.1093/rfs/hhp080.
- Shaffer, S. 1998. The Winner's Curse in Banking. *Journal of Financial Intermediation* 7:359–92.
- Sharpe, S. 1990. Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships. *Journal of Finance* 45:1069–87.
- Stein, J. 2002. Information Production and Capital Allocation: Decentralized versus Hierarchical Firms. *Journal of Finance* 57:1891–921.
- Sussman, O., and J. Zeira. 1995. Banking and Development. CEPR Discussion Paper 1127.
- von Thadden, E.-L. 2004. Asymmetric Information, Bank Lending and Implicit Contracts: The Winner's Curse. *Finance Research Letters* 1:11–23.