

The Information Revolution and Small
Business Lending: The Missing Evidence

Robert DeYoung, W. Scott Frame,
Dennis Glennon, and Peter Nigro

Working Paper 2010-7
March 2010

The Information Revolution and Small Business Lending: The Missing Evidence

Robert DeYoung, W. Scott Frame, Dennis Glennon, and Peter Nigro

Working Paper 2010-7

March 2010

Abstract: This paper provides empirical confirmation for Petersen and Rajan's (2002) widely accepted conjecture that information technology was the primary driver of the observed increase in small business borrower-lender distances in the United States in recent years. Using a different data source for small business loans, we show that annual increases in borrower-lender distances were slow and steady prior to 1993 (the end point in Petersen and Rajan's data) but accelerated rapidly after that. Importantly, we are able to assign at least half of this acceleration to the adoption of credit scoring technologies by the lending banks. Our tests also reveal strong statistical associations between lending distances and borrower characteristics, lender characteristics, market conditions, regulatory constraints, moral hazard incentives, and principal-agent incentives.

JEL classification: G21, O33

Key words: borrower-lender distance, credit scoring, information technology, small business lending

The authors thank Sumit Agarwal and two anonymous referees for comments on an earlier draft, Pam Frisbee for research assistance, and Dan McMillen for his generosity of time and spirit. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta, the Federal Reserve System, the Office of the Comptroller of the Currency, or the U.S. Treasury Department. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Robert DeYoung, University of Kansas, 226-H Summerfield Hall, School of Business, 1300 Sunnyside Avenue, Lawrence, KS 66045-7585, 785-864-1806, rdeyoung@ku.edu; W. Scott Frame, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8783, scott.frame@atl.frb.org; Dennis Glennon, Economics Department, Office of the Comptroller of the Currency, 250 E Street, S.W., Washington, DC 20219-0001, 202-874-4725, dennis.glennon@occ.treas.gov; or Peter Nigro, Finance Department, Bryant University, 1150 Douglas Pike, Smithfield, RI 02917-1285, 401-232-6380, pnigro@bryant.edu.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's Web site at www.frbatlanta.org/pubs/WP/. Use the WebScriber Service at frbatlanta.org to receive e-mail notifications about new papers.

The Information Revolution and Small Business Lending: The Missing Evidence

In one of the more widely cited papers in the recent finance literature, Petersen and Rajan (2002) document small, but systematic, increases in the geographic distances between small businesses and their bank lenders between 1973 and 1993. The authors primarily attribute this increase in borrower-lender distances to the greater use of information technology by lenders, empirically rejecting other potential causal factors such as banking industry consolidation and changes in the distribution of borrower locations over time. They conclude that better, cheaper, and faster access to information, which includes but is not limited to hard information about borrower creditworthiness, allows banks to lend to increasingly more distant firms without compromising their ability to successfully underwrite these credits, monitor the borrowers, and intervene when necessary. This conclusion has become a stylized fact in the financial economics literature and has motivated numerous studies of the impact of increasing distance on credit access, loan pricing, and loan and lender performance (see, e.g., Degryse and Ongena (2005); Agarwal and Hauswald (2006); Berger and DeYoung (2006); Brevoort and Hannan (2006); and DeYoung, Glennon, and Nigro (2008)).

Notably, Petersen and Rajan (2002) draw their key inference mainly from indirect evidence, which contains neither the application of information technology nor differences in lending production functions over time or across banks. Instead, the authors merely show that: (i) borrower-lender distance continued to increase over time after conditioning on other factors, and (ii) the ratio of bank employment-to-bank lending declined between 1973 and 1993, consistent with the substitution of hard information technology inputs for labor-related soft information inputs. Moreover, their database largely pre-dates the implementation of credit scoring models for small business lending (Akhavain, Frame, and White 2005), a strategically transcendent information technology that has transformed the retail banking production process.

In the absence of direct observable evidence, the inference drawn by Petersen and Rajan (2002) about the drivers of borrower-lender distances is an impressive accomplishment. Nonetheless, it is important to demonstrate whether, and to what degree, the inference is correct. We are not aware of any study that has either confirmed or refuted the Petersen and Rajan conjecture based on the observable application of information technology to small business lending. In this short paper, we establish a strong empirical link between small business credit scoring and borrower-lender distance for a sample of 31,880 Small Business Administration (SBA) loans originated between 1984 and 2001. During the first portion of our database, which overlaps the final ten years of the Peterson and Rajan data, we find small but steady annual increases in borrower-lender distances that are very similar to the pattern documented by those authors. But the second portion of our database—which extends beyond their data and coincides with the adoption of small business credit scoring methods by many large U.S. banking companies—yields quantitatively different patterns. We find a dramatic acceleration in small business borrower-lender distance well beyond that documented by Petersen and Rajan, and that the increases in lending distance were substantially larger for loans written by credit-scoring banking companies. Thus, we provide the direct link between information technology and increased lending distances that is absent from the original study. Moreover, by combining our results with those of Petersen and Rajan we are able to document a three decade-long expansion in small business borrower-lender distances, which begins with increased lender reach within local markets and ends with the near-complete erosion of geographic boundaries.

We first establish these findings simply and starkly using a graphical analysis, and we then confirm the findings by estimating controlled multivariate regressions. In addition to establishing point estimates for the impact of credit scoring on lending distance, the regression analysis yields a number of important and interesting ancillary results. We show that small business borrower-lender distance is associated in reasonable ways with borrower characteristics, lender characteristics, market conditions, regulatory constraints, moral hazard incentives (lender exploitation of government guarantees), and principal-agent incentives (borrower exploitation of corporate limited liability). After controlling for

these influences, we are able to assign at least half of the post-1993 (i.e., post-Petersen and Rajan data) acceleration in borrower-lender distances to the adoption of small business credit scoring technologies by the lending banks.

The paper proceeds as follows. Section 1 describes our two main databases: a random sample of loans from the SBA's flagship 7(a) program and the Federal Reserve Bank of Atlanta's database on U.S. commercial bank adoption of small business credit scoring methods. Section 2 analyzes these data at the aggregate level, by comparing the annual trend in average borrower-lender distance to the time series of distance data reported by Petersen and Rajan (2002). Section 3 analyzes the loan-level data by estimating a variety of multivariate regressions. Section 4 concludes the study.

1. The Data

Petersen and Rajan (2002) extracted borrower-lender distances from the Federal Reserve's 1993 National Survey of Small Business Finances (NSSBF). By observing in what year the small businesses included in the survey had first borrowed from their current (i.e., 1993) lender, the authors were able to show that borrower-lender distances at origination had been increasing during the preceding 20 years. In contrast, we directly observe borrower-lender distances for Small Business Administration (SBA) loans originated in each year from 1984 through 2001. The NSSBF data sample the population of all small firms (fewer than 500 employees) while the small firms in our sample are all to some extent credit-constrained borrowers as defined by the SBA—nevertheless, the two samples yield very similar borrower-lender distances during the years in which they overlap, as shown below.¹

The SBA's 7(a) loan program provides loan guarantees to eligible businesses through qualified financial institutions (mainly, but not exclusively commercial banks). Lenders select the firms to receive

¹ The NSSBF does not identify which of its firms participate in SBA loan guarantees; logically, the two samples must overlap because the NSSBF draws its sample from the entire population of U.S. small businesses, a large portion of which have SBA-guaranteed loans. As a rough approximation of this overlap, we note that in 1999 the total dollar value of small business loans under \$250,000 in the U.S. was \$195 billion, while the combined managed SBA guaranteed loan portfolio in 1999 (mostly loans under \$250,000) was just over \$40 billion, or roughly 20 percent of the total (Office of Advocacy, 1999).

loans, initiate SBA involvement, underwrite the loans within SBA program guidelines, and monitor and report back to the SBA the progress of these loans. To qualify for an SBA loan guarantee, a small business borrower must be unable to access the same credit on reasonable terms from non-federally guaranteed sources. If this rule was strictly enforced, then the typical SBA borrower would be relatively more credit-constrained than the typical NSSBF borrower, although in practice the lending bank has substantial discretion in interpreting this language and including the loan in the SBA program (e.g., Temkin and Theodos 2008). The SBA guarantee is a *pro rata* loss sharing arrangement between the SBA and the lender; in the event of a default, the SBA absorbs a fraction of each dollar of loss with the lending bank exposed to the balance.²

For purposes of comparison across databases, it is important to ask whether government guarantees alter the risk-return proposition in SBA loans so much that borrower-lender distance is appreciably less constrained by a lender's ability to observe credit risk. We believe that this is unlikely. First, the SBA guarantee gives lenders a put option that covers only a fraction of any loan losses; so long as SBA lenders incur some non-trivial share of potential loan default costs, they face the same qualitative tradeoffs between risk and return as do non-subsidized lenders. Second, the SBA put option simply reduces risk in a manner that is qualitatively similar to credit risk hedging techniques used by issuers of non-subsidized loans (e.g., portfolio diversification, credit derivatives). Third, geographic distance confers information-gathering frictions (e.g., increased travel costs, less frequent in-person contact), which are independent of borrower risk, on both subsidized and non-subsidized lenders.

Unlike Petersen and Rajan (2002), we are able to observe the implementation of a specific information technology used by banks during our sample period: small business credit-scoring.³ In a telephone survey conducted in January 1998, researchers at the Federal Reserve Bank of Atlanta queried

² The guarantee percentage represents the *pro rata* share of the unpaid principal balance and accrued interest up to 60 days that will be covered by the SBA in the event that a borrower defaults on their obligation. Hence, the loss-sharing arrangement between the lender and the SBA does not include a first-loss position (i.e., it is not a senior-subordinate structure).

³ On pages 2552-2553 Petersen and Rajan state: "...we have been able to characterize the changing environment in which small firms and their lenders operate. We have not directly identified the source of this change..."

the lead (largest) bank in each of the 200 largest U.S. bank holding companies on whether and how they used small business credit scoring (SBCS). Of the banks that responded, 63% reported that they were using SBCS as of or prior to the survey date.⁴ Frame, Srinivasan, and Woosley (2001) provide a complete discussion of the survey questions and results.

We capture the technology adoption decision with the dummy variable *SCORER*, which equals one for any bank affiliated with surveyed banking companies that used SBCS during the year in question, and is loan size-adjusted based on information in the Federal Reserve survey.⁵ We make the reasonable assumption that this underwriting technology passed freely from the lead banks queried in the survey to their affiliated banks. Because only half (99 of 200) of the banking companies contacted responded to the Federal Reserve survey, we exclude from our data all loans from banks affiliated with the non-responders. Because we do not know with certainty whether “scoring banks” credit-scored all, or just a portion, of their small business loan applications (although the authors of the survey lean toward the former), we state our credit-scoring hypothesis carefully: We test whether banks that use credit scoring models have different distance patterns, not whether credit-scored loans have different distance patterns. We are not concerned that the survey included only the lead banks of the largest 200 U.S. banking organizations, because small business credit scoring was almost exclusively a large bank activity prior to 1999 (Frame, Srinivasan, and Woosley, 2001). We also are not concerned that our SBA loan data extends three years beyond the survey date; the distance-technology patterns in the raw data are fully apparent well before 1998.

Our SBA loan data represent a 20% random annual sample of the small business loans originated each year from 1984 through 2001 by U.S. commercial banks under the 7(a) loan program. The sample contains 31,880 small business loans made by 5,497 banks. Due to the growth in the loan program over

⁴ Because the earliest known adoption of SBCS by a bank was in 1992 (Akhavain, Frame, and White 2005), the 1973-1993 Peterson and Rajan database for all intents and purposes precedes this innovation. However, the authors presaged the rapidity of its future adoption: “Conversations with industry experts and evidence in recent studies suggests that, if anything, the trend has accelerated since 1993 (page 2534).”

⁵ The survey asked banks whether they used SBCS for all small business loan applications or only for some loan sizes (i.e., loans less than \$100,000; loans between \$100,000 and \$250,000; or loans between \$250,000 and \$1 million). We use this information to code each lender’s *SCORER* value appropriately for the different size SBA loans that are made.

time, our sample is weighted toward more recent years; 1995 has the most loans with 4,436 and 1985 has the least loans with 574. The SBA data identifies the bank lender (including the branch location where the loan was made) and this allows us to merge the SBA loan data with lender financial statements (Federal Reserve Y-9Cs, FDIC Call Reports) as well as with the Atlanta Federal Reserve credit scoring information. The SBA data also identify the location of the small business borrower and this allows us to merge the SBA loan data with local market structure data (FDIC Summary of Deposits) as well as with state-level macroeconomic data. Descriptive statistics for all of the variables derived from these data are displayed in Table 1 and described in detail below in Section 3.

2. Univariate Trends in Small Business Borrower-Lender Distance

With both borrower location and lender location in-hand, we calculate the straight-line (as the crow flies) geographic distance (*DISTANCE*) between the borrower and the bank office at which the loan was written. We plot the annual median values of *DISTANCE* for each year of our 1984–2001 data in Figure 1. These data are taken from DeYoung, Glennon and Nigro (2008, Table 1). We also plot the multi-year medians of borrower-lender distances for the 1973–79, 1980–89, and 1990–93 time periods reported by Petersen and Rajan (2002, Table I).

During the overlapping 1984-1993 time window, both these time series exhibit relatively short borrower-lender distances that increase gradually over time.⁶ These similarities lend support to the backward-looking method used by Petersen and Rajan (2002) to measure the secular trend in borrower-lender distance. The similarities also show that our SBA data (distance at the time of loan origination) are capturing the same technological evolution in loan underwriting and monitoring found by Petersen and Rajan in the NSSBF data.

⁶ Although the two time trends are very similar, there is a systematic and stable gap of 1½ to 2 miles between them. This gap may indicate a survivorship bias in the backward-looking methodology used by Petersen and Rajan, i.e., if small businesses located closer to their lenders were easier to evaluate and monitor, these lending relationships would be more likely to survive until 1993. The gap may also be caused by the partial SBA loan guarantee, which reduces (but does not eliminate) lender loss exposures, possibly inducing banks to lend to firms at longer distances.

The increasing trend in average borrower-lender distance accelerates in the post-1993 SBA data. Specifically, borrower-lender distance increases markedly for SBA loans made by affiliates of large banking companies that used credit-scoring models to evaluate small business loan applications (*SCORER* = 1); but continues its more gradual secular pace for the loans made by non-scoring banks and their affiliates (*SCORER* = 0). This univariate analysis suggests that a specific technology—small business credit scoring—is the most important driver of the post-1993 increases in borrower-lender distance. However, non-trivial changes in information technology, risk management practices, banking industry regulation, and banking market competition also occurred after 1993 and may have contributed to or dampened the increase in borrower-lender distances. We attempt to identify and disentangle these influences in the multivariate regression tests that follow.

3. Multivariate Analysis

To obtain better measurements of these evolving distance phenomena, we estimate the *DISTANCE-SCORER* relationship in a controlled multivariate setting. We model borrower-lender distance as a function of lending technology, lender characteristics, borrower characteristics, loan terms, and local market conditions. The SBA loan originations in our data occurred over an 18 year time period (1984–2001), but we treat the data as a cross section. All variables are observed in the year in which the loan was originated. We use two different specifications to account for changes in scoring-related and non-scoring-related technological change over time. The more flexible of the two specifications includes a vector of annual dummy variables (*YEAR*) on the right-hand side of the regression equation:

$$\begin{aligned} \ln DISTANCE_i = & \alpha + \beta * SCORER_i + \gamma * YEAR_i \\ & + g(LENDER_i; BORROWER_i; LOAN_i; MARKET_i) + \varepsilon_i \end{aligned} \quad (1)$$

The second specification is less flexible but includes a pair of linear time trends (*TIME1* and *TIME2*) and an interaction term (*TIME2*SCORER*) that allow us to estimate slopes for the various data trends displayed in Figure 1:

$$\ln DISTANCE_i = \alpha + \beta * SCORER_i + \gamma_1 * TIME1 + \gamma_2 * TIME2 + \gamma_3 * TIME2 * SCORER_i$$

$$+ g(\text{LENDER}_i; \text{BORROWER}_i; \text{LOAN}_i; \text{MARKET}_i) + \varepsilon_i \quad (2)$$

In both (1) and (2), i indexes individual loans originated during our sample period 1984–2001 and ε is a random error term assumed to be symmetric with the zero mean.

We specify the dependent variable in natural logs because the distribution of *DISTANCE* is highly skewed by the increase in lending distances late in our sample period; for example, while the mean borrower-lender distance in our sample is almost 64 miles, the median borrower-lender distance is only nine miles. (We add one mile to *DISTANCE* before taking the natural log.) The main test variable *SCORER* is a dummy equal to one for lenders affiliated with organizations in which the lead bank identified itself as a “small business credit scorer” in the Atlanta Fed survey. Some lenders reported that they credit scored only certain sized small business loan applications—e.g., loans under \$100,000; between \$100,000 and \$250,000; between \$250,000 and \$1 million—and we adjust the *SCORER* variable accordingly for each lender. *YEAR* is a vector of fixed-effect dummies based on the loan origination year (1984 is omitted). *TIME1* is a linear time trend over the entire 1984-2001 sample period (*TIME1*=0 in 1984; =17 in 2001) and *TIME2* is a linear time trend over the post-1993 portion of the sample period (*TIME2*=0 in 1984-1993; =8 in 2001).

LENDER, *BORROWER*, *LOAN* and *MARKET* are vectors of control variables included to absorb the influence of lender characteristics, borrower characteristics, loan terms, and local market structure on borrower-lender distances. We define the elements in each of these vectors below. Although these variables are included chiefly as controls, we provide a brief analysis and discussion of their estimated coefficient values.

3.1. Main Regression Results

We estimate equations (1) and (2) using ordinary least squares techniques. Statistical inference is performed using two different sets of robust standard errors, clustered alternatively on either the lending banks or on the geographic markets of the borrowers.⁷ Full regression results are displayed in Table 2.

⁷ For urban borrowers, the local market is defined as the borrower’s Metropolitan Statistical Area (MSA). For borrowers located outside of MSAs, the local market is defined as the borrower’s state.

The estimated coefficients on the *SCORER*, *YEAR* and *TIME* variables are strongly consistent with the patterns displayed in Figure 1.

Both regressions indicate that credit-scoring banks reached out substantially farther than non-scoring banks to make small business loans. Based on the equation (1) estimates, the average loan made by a scoring bank was 52.5 miles more distant than the average loan made by a non-scoring bank, *ceteris paribus*. This result is derived as follows: $52.5 = \exp[\ln(42.3) + 0.8068] - 42.3$, where 42.3 is the post-1993 mean value of *DISTANCE* for non-scoring banks. Based on the equation (2) estimates, scoring banks lent an average 29.7 miles farther than non-scoring banks, derived as follows: $29.7 = \exp[\ln(42.3) + 0.5312] - 42.3$, where $0.5312 = \partial \ln \text{DISTANCE} / \partial \text{SCORER}$ evaluated at the post-1993 mean value of $\text{TIME2} = 3.798$. Both of these point estimates are economically meaningful and statistically significant at the 1% level.

Holding the effects of credit scoring constant, the estimated coefficients on the *YEAR* dummies in equation (1) indicate that statistically significant annual increases in borrower-lender distance began in the early-1990s and quickly became economically substantial. For example, the coefficient on *YEAR1994* translates into an average 9.2 mile increase in mean borrower-lender distance between 1984 (the omitted year) and 1994. This result is derived as follows: $9.2 = \exp[\ln(28.8) + 0.2764] - 28.8$, where 28.8 is the mean value of *DISTANCE* for loans originated in 1984. As we observed in Figure 1, the annual increases in borrower-lender distance accelerated during the 1990s; accordingly, performing a similar calculation based on the *YEAR2001* coefficient yields an average 35.9 mile increase in borrower-lender distance over the entire sample period. These substantial increases in distance are attributable to (a) non-credit scoring-related technological advance and (b) the average effect of economic, regulatory, or industry phenomena that are not specified on the right-hand side of equation (1).

The equation (2) specification allows us to separate the trends in borrower-lender distance into three distinct sources: the average trend in *DISTANCE* from 1984 through 2001 *excluding* the effects of post-1993 technological advance (*TIME1*); the additional post-1993 trend in *DISTANCE* associated with non-scoring-related technological advance (*TIME2*); and the additional post-1993 trend in *DISTANCE*

associated with credit scoring ($TIME2*SCORER$). The annual increases in borrower-lender distance during our sample period are allocated relatively equally across these three sources. Excluding the impact of post-1993 technological advances, borrower-lender distance increased by 3.36% per year on average during our 1984-2001 sample period. This result might be interpreted as an extension of the slow increasing trend in borrower-lender distances identified by Petersen and Rajan (2002) for loans originated during 1973-1993. For non-scoring banks, borrower-lender distance increased by an additional 2.40% per year during the post-1993 portion of our data, arguably the result of general (non-scoring-related) technological advances in lending. For scoring banks, borrower-lender distance increased by an additional 3.01% per year ($0.2194*0.1373$, where the latter number is the mean value of $SCORER$ in the data). Hence, conservatively speaking, at least half of the dramatic acceleration in small business borrower-lender distances during the 1990s is attributable to the implementation of small business credit-scoring.

3.2. Control Variable Results

This section briefly defines the control variables included in (1) and (2) and the implications of their estimated coefficients. All numerical interpretation of these coefficients are based on the results from regression (1) in Table 2.

Lender Characteristics. Banks better situated to manage the risk associated with business lending are more likely to lend to more distant small businesses. $DISTANCE$ is further for large banks ($\ln BANKSIZE$), well-capitalized banks ($EQUITYRATIO$), banks that specialize in business lending ($C\&ILOANRATIO$), and banks with SBA accreditation as “preferred loan providers” ($PLP = 1$). Sensibly, $DISTANCE$ tends to be smaller for more spatially concentrated banks—those that operate proportionately more bank branches ($BRANCHES/ASSETS$) or are part of a multi-bank holding company ($MBHC = 1$). After controlling for bank size and organizational form, we find no evidence that the size of the bank holding company ($\ln BHCSIZE$) influences lending distance. We also find no evidence that the quality of a bank’s overall commercial loan portfolio ($CIPASTDUERATIO$) or certification as a “certified loan provider” ($CLP = 1$) are strong determinants of $\ln DISTANCE$.

Borrower Characteristics. Larger SBA borrowers ($\ln FIRM SIZE$) tend to secure credit at lenders located relatively close by, although this effect is a small one: a 10% increase in firm size is associated with only a 0.4% decrease in distance. Controlling for borrower size, young firms ($NEW FIRM = 1$) tend to be located farther away from their lenders on average, which supports the conventional notion that new businesses must search far and wide for credit. Relative to small businesses organized as proprietorships (the omitted variable) or as partnerships, legally incorporated small businesses ($CORPORATION = 1$) tend to be located closer to their lenders; this result is consistent with lenders' awareness of the potential agency problems associated with limited owner liability. Agricultural, Fishery, and Forestry firms ($SIC CODE_A = 1$) are located especially far from their bank lenders because, unlike other businesses located in rural areas, these firms cannot move their production processes into town. Retail firms ($SIC CODE_G = 1$) are typically located in shopping districts, with banking offices close by.

Loan Characteristics. The data indicate that lenders interpret both large government-guarantee coverage ($GUAR\%$) and poor information about the borrower ($LOWDOC = 1$) as signals of risk: borrower-lender distances are shorter for loans with these characteristics. However, lending distances are longer for loans that exhibit both of these characteristics ($GUAR\% * LOWDOC$), which is consistent with lenders exploiting moral hazard incentives, i.e., shielding lenders from credit risk induces them to lend to more distant borrowers. $\ln DISTANCE$ increases with the size of the loan ($\ln LOAN SIZE$), which indicates that borrowers that seek larger amounts of credit are willing to search farther. Borrowers with 7-year loans ($MATURITY7 = 1$) are located somewhat farther away from their lenders than borrowers with 15-year loans (the omitted maturity variable).

Local Market Characteristics. The spatial characteristics of local lending markets matter. $\ln DISTANCE$ is smaller when the borrower is located in a ZIP Code containing more bank branches ($\ln BRANCH ZIP$) and larger when the borrower is located in a state that imposed restrictions on bank branching at the time of the loan ($BRANCH RESTRICT = 1$). Lending distances are longer in *URBAN* markets, consistent with a rural economic landscape that naturally places both small businesses and their potential lenders close together in small towns. Controlling for these spatial characteristics, there is no

statistical relationship between *lnDISTANCE* and the concentration of commercial bank lenders (*HHI* = Hirschman-Herfindahl Index) in the local market. Borrower-lender distances are larger for firms in economically vibrant markets (*INCOMEGROWTH*), which indicates that these markets can attract non-local lenders.

3.3. Additional Tests

To test the robustness of our main findings, we also estimated equations (1) and (2) for subsamples of the data based on: the size of the small business loans (loans less than \$100,000, loans between \$100,000 and \$250,000, and loans larger than \$250,000); the size of the lending bank (bank assets less than or greater than \$1 billion); and the time period in which the loan was originated (the first half of the sample, the second half of the sample, and the pre-1999 years). All of the main results reported in Table 2 were robust in these subsample tests. Using equation (1), the estimated *SCORER* coefficient ranged between 0.41 and 0.96 and was statistically significant in all of the subsample regressions. The estimated *YEAR* coefficients continued to be statistically non-significant during the 1980s and/or early 1990s, after which they became increasingly positive and statistically significant through 2001. Using equation (2), the estimated *SCORER* coefficient ranged between -1.569 and -0.167 and was generally significant, while the estimated *TIME2*SCORER* coefficient ranged between 0.1417 and 0.5480 and was always statistically significant. (Results are not shown here, but are available upon request.)

All of our regression tests are based on the assumption that the lending technology that a bank uses to evaluate a small business loan application is exogenous to the distance of the applicant to the bank. But it is potentially plausible that lending technology is endogenous to borrower-lender distance—that is, banks that lend to more distant small businesses will be more likely to adopt small business credit scoring techniques in order to reduce the high cost of collecting information on these clients—and we performed a simple lender-level test to rule out this possibility. First, we created the new variable called *NEWSCORER* which was equal to one in the year that a bank adopted small business credit scoring, and equal to zero in all other years and for all other banks. Then, using the post-1993 data only (there was no

credit scoring reported prior to that year), we estimated a binary logit model in which a bank's year t value of *NEWSCORER* is determined by the average value of *DISTANCE* for the loans it originated in year $t-1$. The coefficient on lagged *DISTANCE* was negative and very small (-0.0002) with a p -value of 0.089 for a sample size of 9,829 observations. Thus, we reject the possibility that *SCORER* is endogenous to *DISTANCE*.

4. Conclusions

The distances between small businesses and their bank lenders in the U.S. have been on the increase for several decades. If technological advance has been the catalyst for these increases—as opposed to banking industry consolidation, say, which could lead to increased distances by reducing the number of lenders in any given geographic area—then the phenomenon indicates increased competition and greater supply of credit in local markets. Improved information technology, when applied to credit underwriting and monitoring, allows small businesses to more credibly demonstrate their creditworthiness to lenders and permits out-of-market lenders to more effectively gauge and manage credit risk in more distant local markets.

Petersen and Rajan (2002) observed marginal increases in small business borrower-lender distances during the 1980s and early 1990s, and were the first to infer from the data that these increases were made possible by new hard information technologies. This inference has since become a stylized fact in the banking literature. However, Petersen and Rajan did not directly quantify or analyze the implementation of information technology at these banks, but merely conjectured that this must be the case based on circumstantial evidence—they observed that labor-to-loans ratios were decreasing at lending banks, which is consistent with the substitution of hard information technology for soft information labor inputs. Surprisingly, to date no study has either confirmed or refuted their conjecture.

We make three important findings in this study. First, using a small business loan database different from the one employed by Petersen and Rajan (2002), we corroborate their findings that: (i) borrower-lender distances were increasing during the 1980s and early 1990s; (ii) small businesses were

located quite close to their lenders, typically just a few miles away; and (iii) the rate at which average borrower-lender distances were increasing was slight, measuring only a fraction of a mile each year. Second, we demonstrate that the annual increases in small business borrower-lender distances have in recent years accelerated substantially beyond the distances reported by Petersen and Rajan. Third, and most importantly, we provide direct evidence that links the largest and most recent increases in borrower-lender distance to a specific, hard information-based innovation in information technology: small business credit scoring. Our estimates indicate that the average borrower-lender distance between 1984 and 2001 was nearly 36 miles longer for loans made by banks that used small business credit scoring techniques.

Our results show that small business credit-scoring methods are a primary driver of increased borrower-lender distances in the U.S. Our estimates also suggest that a non-trivial portion of increased borrower-lender distance is attributable to general (i.e., non-scoring) improvements in information technology: using two different regression specifications to capture these trends, we find that borrower-lender distances have also accelerated (although not by as much) for non-scoring banks. This finding both corroborates and extends the increases in borrower-lender distances documented by Petersen and Rajan (2002) during the pre-scoring era prior to 1993. By separating the scoring and non-scoring drivers of distance, our results show that even the most traditional of lending models—that is, lending to opaque business borrowers based on soft information and bank-borrower relationships—are made more efficient by advances in information, communications, and other technologies.

Consequently, our tests generate substantial evidence that phenomena other than information technology—such as cross-sectional differences in borrower, lender, and market characteristics—are important determinants of small business borrower-lender distances. One of our more intriguing findings: larger government loan guarantees make lenders more willing to lend to especially information-poor borrowers (i.e., those applying for low-doc loans) at longer distances. This result is consistent with lenders exploiting moral hazard incentives—a taxpayer shield from credit losses induces lenders to adopt riskier lending practices.

References

- Agarwal, Sumit, and Robert Hauswald, 2006, Distance and Information Asymmetries in Lending Decisions, Working paper, American University.
- Akhavein, Jalal, W. Scott Frame, and Lawrence J. White, 2005, The Diffusion of Financial Innovation: An Examination of the Adoption of Small Business Credit Scoring by Large Banking Organizations, *Journal of Business* 78(2), 577-96.
- Berger, Allen N., and Robert DeYoung, 2006, Technological Progress and the Geographic Expansion of Commercial Banks, *Journal of Money, Credit, and Banking* 38(6), 1483-1513.
- Brevoort, Kenneth and Timothy H. Hannan. 2006. "Commercial Lending and Distance: Evidence from Community Reinvestment Act Data." *Journal of Money, Credit, and Banking*, 38(8), 1991-2012.
- Cole, Rebel A. and John D. Wolken. 1995. "Financial Services Used by Small Businesses: Evidence from the 1993 National Survey of Small Business Finances," *Federal Reserve Bulletin*, July, 659-667.
- DeYoung, Robert, Dennis Glennon, and Peter Nigro. 2008. "Borrower-Lender Distance, Credit Scoring, and the Performance of Small Business Loans." *Journal of Financial Intermediation*, 17(1), 113-143.
- Frame, W. Scott, Aruna Srinivasan, and Lynn Woosley. 2001. "The Effect of Credit Scoring on Small Business Lending." *Journal of Money, Credit, and Banking*, 33(3), 813-825.
- Office of Advocacy, U.S. Small Business Administration, *Small Business Lending in the United States, 1999 Edition*.
- Petersen, Mitchell A., and Raghuram G. Rajan. 2002. "The Information Revolution and Small Business Lending: Does Distance Still Matter?" *Journal of Finance*, 57, 2533-2570.

Table 1
Summary Statistics

Variables used in the estimation of equations (1) and (2). Data are a random sample of 31,880 Small Business Administration 7(a) loans originated by U.S. commercial banks between 1984 and 2001. All variables are defined in the text.

Variable	Mean	Standard Deviation	Minimum	Maximum
DISTANCE	63.6516	286.6681	0.1000	7882.6000
lnDISTANCE	2.2000	1.9258	-2.3026	8.9724
SCORER	0.1373	0.3441	0	1
Preferred Lender Program	0.1283	0.3344	0	1
Certified Lender Program	0.1629	0.3693	0	1
BANK SIZE	\$14,601,626.63	60,954,976.51	\$1,055.93	\$607,085,000
lnBANKSIZE	13.1973	2.4575	6.9622	20.2242
BHC SIZE	\$51,035,601	154,737,291	\$1,6247	\$1.26E+09
lnBHCSIZE	13.9822	2.9064	7.301148	20.38828
BRANCHES/ASSETS	0.0690	0.1761	0.0015	7.5759
LOGBRANCHZIP	1.7116	0.884829	0	4.189655
MBHC	0.5316	0.499007	0	1
EQUITYRATIO	8.3816	2.9972	-5.59	76.83
CILOANRATIO	0.1576	0.0910	0	0.741675
CIPASTDUERATIO	2.2728	10.7817	0	1195.83
HHI	0.1906	0.1150	0	1
URBAN	0.8232	0.3815	0	1
INCOMEGROWTH	1.2371	0.9171	-5.76254	8.182991
BRANCHRESTRICT	0.8339	0.3722	0	1
CORPORATION	0.5605	0.4963	0	1
PARTNERSHIP	0.0735	0.2609	0	1
FIRMSIZE*	12.5077	111.4224	1	9999
lnFIRMSIZE	1.6533	1.1279	0	9.2102
NEWFIRM	0.3268	0.4691	0	1
Ag., Fishery, Forestry (SIC_A)	0.0321	0.1762	0	1
Mining (SIC_B)	0.0034	0.0581	0	1
Retail (SIC_G)	0.3321	0.4710	0	1
LOANSIZE	\$144,242	165,037	\$2,000	\$2,550,000
lnLOANSIZE	11.4240	0.9505	7.600903	14.7516
GUAR%	0.8003	0.1011	0.11	0.9
LOWDOC	0.3108	0.4628	0	1
MATURITY3	0.1376	0.3445	0	1
MATURITY7	0.6455	0.4784	0	1

* The SBA 7(a) loan program has a firm-size constraint of 500 employees. However, the SBA can and does occasionally approve loans to firms that exceed this size threshold. Of the 31,880 loans in our sample, 15 borrowing firms had more than 500 employees.

Table 2
Borrower-Lender Distance Regressions

Results from the ordinary least squares estimation of equations (1) and (2) for a random sample of 31,880 Small Business Administration 7(a) loans originated by U.S. commercial banks between 1984 and 2001. The dependent variable *lnDISTANCE* is the natural log of borrower-lender distance. *SCORER* is a dummy equal to one if the lending bank was affiliated with an organization in which the lead bank used credit scores to evaluate small business loan applications. *TIME1* is a linear time trend over the entire 1984-2001 sample period. *TIME2* is a linear time trend over the post-credit scoring 1993-2001 portion of the sample period. The *YEAR* variables are dummies indicating the year of loan origination (1984 is omitted). All other variables are defined in the text.

Variable	(1)			(2)		
	coefficient	<i>p-values</i>		coefficient	<i>p-values</i>	
		clustered on lender	clustered on market		clustered on lender	clustered on market
<i>Intercept</i>	-0.9734	0.074	0.036	-1.3025	0.0228	0.0057
<i>SCORER</i>	0.8068	0.000	0.000	-0.3021	0.2543	0.055
<i>TIME1</i>				0.0336	0.0001	0.0001
<i>TIME2</i>				0.0240	0.0872	0.0371
<i>TIME2*SCORER</i>				0.0240	0.0001	0.0001
<i>YEAR1985</i>	-0.0096	0.928	0.929			
<i>YEAR1986</i>	-0.1406	0.126	0.117			
<i>YEAR1987</i>	0.0910	0.356	0.297			
<i>YEAR1988</i>	0.0570	0.564	0.554			
<i>YEAR1989</i>	0.0307	0.733	0.726			
<i>YEAR1990</i>	0.0521	0.566	0.559			
<i>YEAR1991</i>	0.2360	0.014	0.010			
<i>YEAR1992</i>	0.1363	0.147	0.118			
<i>YEAR1993</i>	0.1665	0.058	0.052			
<i>YEAR1994</i>	0.2764	0.002	0.001			
<i>YEAR1995</i>	0.2873	0.001	0.001			
<i>YEAR1996</i>	0.4383	0.000	0.000			
<i>YEAR1997</i>	0.4227	0.000	0.000			
<i>YEAR1998</i>	0.5655	0.000	0.000			
<i>YEAR1999</i>	0.5992	0.000	0.000			
<i>YEAR2000</i>	0.7029	0.000	0.000			
<i>YEAR2001</i>	0.8093	0.000	0.000			
<i>Lender Characteristics:</i>						
<i>Preferred Lender Program</i>	0.3960	0.000	0.000	0.4284	0.0001	0.0001
<i>Certified Lender Program</i>	0.0542	0.169	0.202	0.0653	0.0960	0.1237
<i>lnBHCSIZE</i>	-0.0002	0.992	0.987	-0.0013	0.9566	0.9305
<i>lnBANKSIZE</i>	0.1594	0.000	0.000	0.1548	0.0002	0.0001
<i>BRANCHES/ASSETS</i>	-0.2143	0.332	0.264	-0.2346	0.3345	0.2539
<i>MBHC</i>	-0.0829	0.068	0.020	-0.0792	0.0789	0.0269
<i>EQUITYRATIO</i>	0.0193	0.019	0.001	0.0188	0.0191	0.0007
<i>CILOANRATIO</i>	0.4475	0.044	0.017	0.4784	0.0332	0.0082
<i>CIPASTDUERATIO</i>	0.0011	0.092	0.124	0.0010	0.0976	0.1316

Table is continued on next page

Table 2--continued

Variable	(1)			(2)		
	coefficient	<i>p-values</i>		coefficient	<i>p-values</i>	
		clustered on lender	clustered on market		clustered on lender	clustered on market
Borrower Characteristics:						
<i>CORPORATION</i>	-0.0740	0.006	0.013	-0.0728	0.0066	0.0135
<i>PARTNERSHIP</i>	-0.0233	0.613	0.568	-0.0244	0.5948	0.5486
<i>lnFIRMSIZE</i>	-0.0395	0.004	0.005	-0.0373	0.0081	0.0094
<i>NEWFIRM</i>	0.0620	0.011	0.009	0.0637	0.0089	0.0063
<i>Ag., Fish, Forest (SIC_A)</i>	0.3316	0.000	0.000	0.3504	0.0001	0.0001
<i>Mining (SIC_B)</i>	0.1252	0.542	0.390	0.1526	0.4618	0.3054
<i>Retail (SIC_G)</i>	-0.0605	0.008	0.003	-0.0580	0.0129	0.0055
Market Characteristics:						
<i>lnBRANCHZIP</i>	-0.2532	0.000	0.000	-0.2523	0.0001	0.0001
<i>HHI</i>	-0.0026	0.991	0.992	-0.0073	0.9759	0.9766
<i>URBAN</i>	0.4771	0.000	0.000	0.4819	0.0001	0.0001
<i>HHI*URBAN</i>	0.0747	0.825	0.848	0.0805	0.8112	0.8355
<i>INCOMEGROWTH</i>	0.0220	0.079	0.078	0.0230	0.0615	0.0664
<i>BRANCHRESTRICT</i>	0.1140	0.039	0.065	0.1226	0.0224	0.0426
Loan Characteristics:						
<i>lnLOANSIZE</i>	0.1155	0.000	0.000	0.1228	0.0001	0.0001
<i>GUAR%</i>	-1.1677	0.003	0.001	-0.9614	0.021	0.0053
<i>LOWDOC</i>	-3.2305	0.000	0.000	-2.6973	0.0001	0.0001
<i>GUAR%*LOWDOC</i>	3.7599	0.000	0.000	3.1685	0.0001	0.0001
<i>MATURITY3</i>	-0.0704	0.400	0.304	-0.0739	0.3759	0.295
<i>MATURITY7</i>	0.0555	0.108	0.096	0.0551	0.1129	0.0978
<i>Adjusted R-square</i>	0.2071			0.2105		

Figure 1

This figure compares median annual borrower-lender distances constructed for small business loans from two different times and databases. Petersen and Rajan (2002, Table I, page 2,537) constructed borrower-lender distances for banking relationships that started in 1973 and continued through 2001, which are based on data from the 1993 National Survey of Small Business Finances (NSSBF). We construct borrower-lender distances for Small Business Administration (SBA) loans originated between 1984 and 2001. The SBA loans are also reported separately after 1993 for banking organizations that did or did not use small business credit scoring, which are based on data from a Federal Reserve Bank of Atlanta survey (Frame, Srinivasan, and Woosley 2001).

