Lender-Borrower Relationships and Loan Origination Costs

Philip Ostromogolsky*
Center for Financial Research
Federal Deposit Insurance Corporation

January 2017

Preliminary and incomplete. Please do not cite without permission.

Abstract

Using a recently developed method of causal inference, this paper estimates the additional up-front loan origination costs that a small business can expect to pay when it first borrows from a new lender. I compare firms that borrow from a previously unused financial institution with firms that borrow from a financial institution with which they have a preexisting financial relationship.† I estimate that firms that borrow from a new financial institution can expect to pay $5,650 to $6,980 more in closing costs than firms that return to a previously-used financial institution. Based on these findings, I argue that a central function of origination fees is to pay for the production of detailed, firm-specific information that is valuable to the lender. I study a natural quasi-experiment wherein, for a small group of firms, selection into borrowing from a new lender is close to random. Returning to the wider population of small business borrowers, I use the method of Altonji, Elder, and Taber (2002, 2005) to account for endogeneity in firms' selection to borrow from a new lender. The method of Altoji, Elder, and Taber allows me to measure the degree to which a firm's selection to borrow from a new lender is driven by unobservables that also determine closing costs and to correct for any resulting bias. All analyses confirm that borrowing from a new financial institution causes firms to pay higher loan origination costs.

*The analysis, opinions, and conclusions presented here are those of the author alone and do not necessarily reflect the views of the Federal Deposit Insurance Corporation. I am grateful to Matthew Spiegel for invaluable guidance and support. I thank Rosalind Bennett, Gary Gorton, Levent Guntay, Vivian Hwa, Emily Johnston-Ross, Stefan Jacewitz, Pavel Kapinos, Stephen Karolyi, Troy Kravitz, Paul Kupiec, Myron Kwast, Yan Lee, Stefan Lewellen, Steven Malliaris, Andrew Metrick, Oscar Mitnik, Justin Murfin, Jon Pogach, Carlos Ramirez, Jack Reidhill, Boudhayan Sen, Shyam Sunder, Heather Tookes, Haluk Unal, and Smith Williams for helpful comments and much-needed suggestions. Author can be contacted at postromogolsky@fdic.gov.

†A relationship between a firm and a financial institution is said to exist if the firm has previously conducted business with the financial institution.

JEL Classification: G20, G21, G32, L26, C31.

Keywords: small business finance, small business lending, small business loans, lending relationships, loan contracts, debt contracts, switching costs, information costs, causal inference, quasi-experiment.

Contact Information

Philip Ostromogolsky
Center for Financial Research
Federal Deposit Insurance Corporation
550 17th Street NW
Washington, DC 20429

1-202-898-6585
postromogolsky@fdic.gov

1. Introduction

Many small business borrowers form exclusive, long-term relationships with their lenders.[1] Several studies have documented the costs and benefits of bank-firm relationships for small business borrowers. Studies find that firms that maintain longer, stronger relationships with their lenders have more access to credit, pledge less collateral, and pay lower interest rates.[2] This paper focuses on lender-borrower relationships and a previously little-studied cost of small business borrowing — loan origination costs.[3]

Using the method of Altonji, Elder, and Taber (2002, 2005) I estimate the causal effect of borrowing from a new lender on the closing costs that a small business pays at loan origination. I compare firms that borrow from a previously-unused financial institution with firms that borrow from an institution with which they have a preexisting relationship.[4] A simple comparison of means and elementary OLS show that small businesses that turn to a new financial institution pay $5,650 to $6,740 more in closing costs than firms that return to a previously-used institution. The estimated causal effect of borrowing from a new bank is almost identical at $5,740 to $6,980. I also study a natural quasi-experiment wherein, for a subset of firms, whether a firm borrows from a previously-used bank or a new bank is close to randomly determined. Estimates from the natural experiment show that, among this subset of firms, the effect of borrowing from a new bank is $10,140 to $13,230. Based on these findings, I argue that a central function of origination fees is to pay for the production of detailed, firm-specific information that is valuable to the lender.

To begin, I present regression estimates showing that firms that borrow from a previously-unused financial institution pay higher loan origination costs than firms who borrow from an institution which they had used in the past. The regressions contain a host of detailed controls. Furthermore, I exclude data wherein selection into using a new financial institution is likely driven by factors that also determine loan origination costs. The regression analysis provides strong evidence that borrowing from a new lender has a positive and significant causal effect on loan origination costs.

Banks screen small business borrowers for credit quality and monitor borrowers to ensure repayment. Banks form long-term relationships with their borrowers to obtain private, borrower-specific information that is then used to more accurately monitor borrowers and more precisely gauge borrower credit quality.[5] This paper's central finding is that having a preexisting relationship with its lender reduces a firm's loan origination costs (henceforth termed *the relationship effect*). Given the informational value of lender-borrower relationships, this finding provides strong evidence that loan origination fees pay for the production of information that a lender would have otherwise obtained over

[1] Among U.S. small businesses, the reason most frequently cited by a firm for its choice of lender is the existence of a prior relationship with the lender. The mean small business has loans outstanding from 1.02 institutions, and the median small business has loans outstanding from 1 institution (2003 Survey of Small Business Finances, federalreserve.gov/pubs/oss/oss3/nssbftoc.htm).

[2] See Petersen and Rajan (1994), Berger and Udell (1995), Cole (1998), Hellmann, Lindsey, and Puri (2008), Jiangli, Unal, and Yom (2008), and Bharath et al. (2011).

[3] To the best of the author's knowledge, this is the first academic article to study small business loan origination costs.

[4] A relationship between a firm and a financial institution is said to exist if the firm has previously conducted business with the financial institution.

[5] See Freixas and Rochet (2008), Boot (2000), and DeGryse, Kim, and Ongena (2009).

the course of a lender-borrower relationship. My findings suggest that the information produced during loan origination is similar to the information gleaned from a long-standing bank-firm relationship.

At the same time, the relationship effect could be the result of banks offering customers loyalty discounts (similar to frequent-flyer miles), wherein the purchase of a banking service (such as a line of credit) allows a customer to obtain a discount on loan origination fees in the future. I argue that the relationship effect is not caused by repeat-use discounts (written or unwritten). Loyalty programs aim to affect customer purchasing behavior by relying on customers' sensitivity to prices.[6] Loyalty programs offer discounts on pricing terms that are most salient to customers; frequent-flyer programs offer discounts on ticket prices not on booking and luggage fees. I show that, when selecting a lender, firms are much more concerned with loan interest rates (and other loan contract terms) than with loan origination costs. Meanwhile, firms that borrow from a previously-used lender do not see lower loan interest rates. It is unlikely that banks try to build customer loyalty by offering discounts on the least salient pricing terms.[7]

I use two distinct approaches to estimate the causal effect of borrowing from a new financial institution on loan origination costs: (1) a natural quasi-experiment and (2) the econometric method of Altonji, Elder, and Taber (2002, 2005). First, I study a natural quasi-experiment wherein selection to borrow from a new lender is close to randomly assigned. I identify a small group of firms that select a lender based on the lender's physical proximity to the firm. Firms have no control over the opening or closing of financial institutions and their branches. For a firm that chooses a lender based on proximity, lender selection is driven by financial institutions' unpredictable entry and exit into the firm's local market. Consequently, whether a distance-minimizing firm borrows from a previously-used lender or a new lender is uncorrelated with firm characteristics. Data on firms that select a lender based on proximity reveal that borrowing from a new lender significantly raises a firm's loan origination costs.

While the natural experiment described above is informative, the small group of firms it studies may not be representative of the wider population of U.S. small businesses. To estimate the causal effect of borrowing from a new lender among the general population of firms in my data I turn to the novel econometric method of Altonji, Elder, and Taber (henceforth, AET). I implement the method of AET to identify a lower bound on the size of the causal effect of borrowing from a new lender. The method posits a linear causal model with a large number of observable controls and a large number of unobservable variables (aggregated in the error term). The method makes the identifying assumption that the observable controls are, as a group, representative of (similar to) the unobservables. Using the observables as a guide to the unobservables, the technique corrects a naïve regression coefficient for omitted variable bias.

The method of AET reveals that the minimum plausible causal effect of borrowing from a new lender is still positive and significant. Indeed, I find that the true causal effect is most likely greater than the estimate obtained from simple regression. To the

[6] See Klemperer (1987) and Sharp and Sharp (1997).

[7] This argument generalizes to any discount designed to affect customer behavior, such as a "thank you" discount meant to elicit reciprocity. Discounts must be material. Attaching discounts exclusively to the least material pricing terms would be a poor marketing strategy.

best of my knowledge, this paper is the first article in the academic finance literature to implement the method of Altonji, Elder, and Taber.

This paper is motivated by the following hypothesis: Borrowing from a new lender should have a positive causal effect on loan origination costs. At the time of loan application and during negotiations over the loan contract, a prospective lender asks the borrower to produce legal and financial documents that provide critical information about the firm. The lender uses this information to verify the borrower's credit worthiness and to evaluate the borrower's collateral and operations. Frequently, the borrower must pay outside inspectors and appraisers to provide independent assessments of the firm's assets and collateral. Additionally, the borrower often retains legal counsel to help structure the loan contract.

Loan origination costs also include fees charged directly by the lender to the firm. Banks charge loan application fees, origination fees, and legal fees to defray the costs of loan processing. Also, borrowers often pay discount points — a lump-sum payment equal to a percentage of the principal — in exchange for a lower interest rate.

When a bank considers the loan application of a new customer, the bank needs to conduct thorough due diligence to evaluate the prospective borrower's credit worthiness. The loan applicant will be asked to produce a large volume of costly information for the bank to review. In contrast, when a bank considers a customer, with which it has a long-standing relationship, the lender's past experience with the firm is likely to obviate the need for costly due diligence. The lender's knowledge of the firm's assets and business practices, will likely allow the firm to forego the production of costly information.

This paper adds to the growing literature on the benefits of lender-borrower relationship for small business borrowers (Petersen and Rajan, 1994; Berger and Udell, 1995; Cole, 1998).[8] It also adds to the literature on the benefits of bank-firm relationships for large corporate borrowers (Hellmann, Lindsey, and Puri, 2008; Jiangli, Unal, and Yom, 2008; Berg, Saunders, and Steffen, 2013; Bharath et al., 2011). More broadly, this article underscores the importance of long-term relationships in financial intermediation (Puri, Rocholl, and Steffen, 2011; Puri, Rocholl, and Steffen, 2013; Iyer and Puri, 2012; Iyer, Puri, and Ryan, 2015). This paper highlights the importance of fees in loan pricing, complementing the work of Berg, Saunders, and Steffen (2013, 2015), who present the first comprehensive studies of loan fees in large syndicated loans.

From a methodological perspective, this paper adds to the increasing number of banking research papers that use natural experiments (Berg, Puri, and Rocholl, 2014; Agarwal and Wang 2009; Krishnan, Nandy, and Puri; Iyer Peydro, 2011; Ziebarth, 2013). This study also adds to the growing literature on causal inference using observables as a guide to the unobservables (Altonji, Elder, and Taber, 2008; Oster, 2014).

[8] See Section 2 for a review of the relevant literature. For a more detailed survey of the research on lender-borrower relationships see Freixas and Rochet (2008), DeGryse, Kim, and Ongena (2009), Boot (2000), Ongena and Smith (2000), Petersen (1999), and Udell (2008).

2. The Literature on Loans and Financial Relationships

One of the first papers to study small business lender-borrower relationships was by Petersen and Rajan (1994). The authors show that firms with longer lender-borrower relationships are less credit constrained (have a greater availability of credit). Similarly, Cole (1998) finds that a lender is more likely to extend credit to a small business if the lender and the firm have a preexisting relationship. In contrast to Petersen and Rajan, Cole finds that the length of the relationship does not matter. Looking at the intensity of a bank-firm relationship — the number of lenders used by a firm — Petersen and Rajan (1994) find that small businesses that borrow from fewer institutions are less likely to be credit constrained. Similarly, Jiangli, Unal, and Yom (2008) find that during the Asian financial crisis Korean and Thai firms that had more concentrated lender-borrower relationships were more likely to obtain credit than similar firms with less concentrated relationships.

In addition to better access to credit, studies have found that lender-borrower relationships can reduce the cost of borrowing. Looking at lines of credit, Berger and Udell (1995) find that small business borrowers with longer bank-firm relationships pay lower interest rates. Hellmann, Lindsey, and Puri (2008) find that firms that borrow from banks with which they have a preexisting relationship pay lower interest rates than firms that borrow from banks with which they have no prior relationship. Looking at syndicated loans obtained by U.S. corporations, Bharath et al. (2011) find that repeated borrowing from the same lender is associated with a reduction in loan spreads. Berg, Saunders, and Steffen (2013), discussed below, present a similar finding. Uzzi (1999) and Brick and Palia (2007) also find that bank-firm relationships reduce the cost of borrowing. Ioannidou and Ongena (2010) present a more nuanced finding. They show that in the first 1.5 years after a firm switches banks, the firm's loan interest rate declines. Then, 1.5 years after the switch, its interest rates start to rise.

In contrast to the just cited studies, DeGryse and Van Cayseele (2000) find that loan interest rates increase with the duration of bank-firm relationships. Petersen and Rajan (1994) find that the length of a lender-borrower relationship has no effect on small business loan interest rates. Elsas and Krahnen (1998) also find that the price of credit is unrelated to the duration of a bank-firm relationship.

Additionally, research suggests that lender-borrower relationships affect loan collateral requirements. Berger and Udell (1995) show that small business borrowers with longer bank-firm relationships are less likely to pledge collateral on their lines of credit. Bharath et al. (2011) find that, for larger corporate borrowers, a preexisting relationship with a lender reduces a firm's collateral requirements. Using historical data on nineteenth century firms, Bodenhorn (2003) presents a similar finding.

Research on financial relationships extends beyond the study of corporate borrowing. Puri, Rocholl, and Steffen (2011) look at retail loan applicants at German banks affected by the U.S. financial crisis of 2008. They find that a bank is less likely to reject loan applicants who have a preexisting relationship with the bank. In a 2013 paper, the same authors show that retail borrowers who have a relationship with their bank prior to applying for a loan default significantly less than borrowers who do not have a prior relationship (Puri, Rocholl, and Steffen, 2013). Agarwal et al. (2009) also find that retail borrowers who have a prior relationship with their lender at the time of loan application are less likely to default than other borrowers. Iyer and Puri (2012) and Iyer, Puri, and

Ryan (2015) analyze long-term relationships between depositors and their banks. In both studies the authors find that depositors who have a long-standing relationship with a bank are less likely to run on the bank when the bank faces a run.

To the best of my knowledge, the present article is the first academic paper to study the connection between small business loan origination costs and lender-borrower relationships. However, this paper is not the first academic study of loan fees and lender-borrower relationships. In a 2013 working paper, Berg, Saunders, and Steffen study loan fees charged on large, syndicated corporate loans. The authors find that if a firm's current lender provided the firm with a loan in the previous five years, the firm will be charged a lower origination fee, a lower letter of credit fee, and a lower drawdown fee (loan interest rate) on its current loan. In a rich follow-up analysis, Berg, Saunders, and Steffen (2015) study 12 different syndicated loan fees. The authors show that (1) fees are used to price options embedded in loan contracts and (2) that fees are also used to screen borrowers based on the likelihood of exercising these options.

3. Definitions and Data

3.1 Definitions

3.1.1 Loan Origination Costs

Loan origination costs are defined as the total dollar amount of fees paid by a borrower to apply for and obtain a loan at the time of loan origination. Fees counted as part of origination costs include application fees, origination fees, points, lawyer fees, appraisals, inspection fees, title transfer fees, environmental survey fees, and other expenses incurred at loan origination.[9]

3.1.2 Borrowing From a New Lender

A firm is said to borrow from a *new financial institution* or a *new lender* if the firm borrows from an institution with which it has no prior relationship. A firm is said to borrow from an *old lender* or an *old financial institution* if the firm borrows from an institution with which it has a preexisting relationship. A relationship between a small business and a financial institution is said to exist if the firm has previously conducted business with the financial institution. I use the term *previously used* as a synonym for old; I use the term *previously unused* as a synonym for new.

For a given firm, a financial institution is designated as *old* if it has previously provided the firm with a loan. An institution is also designated as *old* if it has not provided the firm with a loan but has provided the firm with some other financial service. Non-loan financial services include (but are not limited to): business checking and savings accounts, credit card processing, transactions services, cash management services, credit related services, and trust and brokerage accounts.[10]

[9] See the *2003 Survey of Small Business Finances Technical Codebook* and the *2003 Survey of Small Business Finances Survey Questionnaire* (www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html).

[10] Ibid.

3.2 Data

3.2.1 Overview of Data

I use data from the 2003 Survey of Small Business Finances (SSBF).[11] The SSBF is a pentennial survey of U.S. firms having fewer than 500 employees conducted by the Board of Governors of the Federal Reserve. The survey targets the entire U.S. population of nonfinancial, nonfarm, nonsubsidiary, business enterprises that were in operation as of December 2003. For each of the 4,240 firms surveyed, the SSBF provides an extremely detailed snapshot of the firm's financial state and financial history. For 1,761 of the surveyed firms, the 2003 SSBF reports information on firms' most recent loan, detailing the loan's origination costs and other vital loan characteristics. For each firm with a most-recent loan, the SSBF records the length of the relationship between the firm and the provider of its most-recent loan at the time of loan origination. At the same time, the data are imperfect, and some observations have missing variables. This paper relies on the 2003 SSBF because the SSBF was discontinued following the publication of the 2003 survey. Thus, the 2003 survey provides the most recent SSBF data available. The SSBF data are supplemented with data on key interest rates taken from the Federal Reserve. All values are measured in 2004 U.S. dollars.[12]

The present analysis focuses on financial relationships between U.S. small businesses and financial institutions. I restrict my analysis to firms that borrow from a financial institution, where a financial institution is one of the following: a commercial bank, a savings bank, a savings and loan association (S&L), a credit union, a finance company, an insurance company, a brokerage firm, a leasing company, or a mortgage company. More than 88 percent of the loans in my data are provided by a commercial or savings bank. I focus my analysis on firms that have access to financial institutions. I restrict my analysis to firms that have at least one relationship with a financial institution at the time of the issuance of their most recent loan. To prevent my analysis of financial relationships from being contaminated by non-financial, social relationships I exclude firms that selected their lender on the basis of a referral, a personal relationship, or family connections. My final dataset consists of 1,261 loan-level observations.

3.2.2 Estimation Framework

My analysis focuses on firms' most recent loan application as of the date of their participation in the 2003 SSBF. In all estimations I model the relationship between the total origination costs paid by a firm on its most-recent loan and a binary indicator variable that equals 1 if the loan is obtained from an institution with which the firm had no business relationship at the time of the firm's loan application. The general form of the estimation equation is given by

$$Cost_i = \kappa + \beta NoRelationship_i + X_i'\gamma + \varepsilon_i \tag{1}$$

[11] See www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html

[12] The 2003 SSBF was initiated in 2003. The majority of firms surveyed by the 2003 SSBF were surveyed between 2004 and 2005. Among the surveyed firms who reported a having a recent loan, the majority applied for the loan in 2004. Consequently, all values are reported in 2004 dollars.

where $Cost_i$ is the total origination cost of the firm's most recent loan (in thousands) and $NoRelationship_i$ is the binary no-relationship indicator. Kappa (κ) is a constant term, and ε_i is the residual. Additionally, X_{i} is a vector of controls and γ is a vector of coefficients. Each observation i contains information on a single loan issued to a single firm. In my analysis *Cost* is the outcome variable of interest, and *NoRelationship* is the treatment indicator variable. For brevity, I will often omit subscripts (as in the sentence above).

The SSBF is a rich dataset, and my estimations use a large set of controls. Table 1 presents a list of all of the controls used. Table 1 also provides definitions for all variables.[13] The controls include the prime rate, default spread, and term spread at the time of loan issuance; firm age and employment; firm assets, leverage, and profits; and a firm's Dun & Bradstreet credit score rank. D&B credit score rank indicates a firm's credit quality relative to other firms evaluated by Dun & Bradstreet. D&B credit score rank is measured on a scale from 0 to 100, with a rank of 100 indicating the highest credit quality a firm can attain. The SSBF does not identify a firm's exact credit score rank. The survey only states that a firm's score lies within one of 6 specific intervals. In order of ascending credit quality, the intervals are: [0,10]; [11,25]; [26,50]; [51,75]; [76,90]; and [91,100]. I use a set of five D&B credit score rank dummies (one is omitted) to control for a firm's credit quality. The regression also controls for the amount of credit granted, and for the presence of loan collateral. The regression includes 2-digit SIC industry fixed effects and census division fixed effects.[14] Also, the regression includes binary indicators for each one of 42 possible lender attributes that a firm can list as a reason for applying to the provider of its most-recent loan. Each firm lists up to three reasons and can have up to three non-zero indicators. Table 1 lists all the reasons. See Table 1 for a list of the controls not discussed above.

3.2.3 Descriptive Statistics

Table 1 presents definitions and summary statistics for all variables. The table shows that the average firm is 18 years old and has $4.9 million in assets and 58 employees. The distributions of firm age, assets, and employment are skewed to the right, with the median firm being 16 years old, having $1 million in assets, and employing 27 people. Most firms are corporations and are located in a Metropolitan Statistical Area (MSA). Table 1 shows that origination costs vary considerably — from zero to $66,000. Thirty-nine (39) percent of recorded origination costs equal zero, while 61 percent are greater than zero. The average loan has an origination cost of $3,160. Among loans with positive origination costs, the mean origination cost equals $5,168.

Table 1 reveals that the vast majority of firms borrow from a lender with which they have a preexisting relationship. Of the 1,261 firms in my sample, 1,177 firms (93.3%) borrow from a previously-used financial institution, and 84 firms (6.7%) borrow from a new lender. The average firm borrows from an institution with which it has a 10-year-

[13] For details on all the variables used in the present analysis see the *2003 Survey of Small Business Finances Technical Codebook* and the *2003 Survey of Small Business Finances Survey Questionnaire* (www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html).

[14] A census division is a major geographic unit of the U.S. The U.S. has nine census divisions. For example, the Northeast Census Division includes the entire northeast U.S. for a total of nine states — from Pennsylvania in the south to Maine in the north.

long relationship. This statistic is driven by firms' tendency to build relationships with the providers of their financial services and firms' inclination to concentrate financial transactions at a single institution.[15] In contrast, among firms that are two years old or younger, 16 percent borrow from a new lender.

Table 2 presents descriptive statistics broken down by *NoRelationship*. The table compares firms that use a lender with which they have a prior relationship (second column) with firms that use a new lender (third column). There is little difference in age and number of employees between firms that have *NoRelationship*=0 and firms that have *NoRelationship*=1. Similarly, firm profitability, cash position, and leverage do not significantly differ by *NoRelationship*. At the same time, firms that borrow from a previously-used lender are on average 5.6 years older than firms that use a new lender. Similarly, the owners of firms that have *NoRelationship*=0 are on average more experienced than the owners of businesses that have *NoRelationship*=1. Interestingly, businesses that borrow from a new lender are located 126 miles farther from their lender than businesses that turn to a previously-used lender. This contrast in *Distance from Firm to Lender* suggests that firms tend to develop relationships with local banks and that firms who borrow from a new lender are willing to or need to shop around within a wider radius.

4. Regression, Causality, and Endogenous Selection

4.1 Regression With Select Controls

I begin my analysis by estimating equation 1 using simple OLS regression and a small set of controls. The results are presented in Table 3. With no controls, the difference in mean origination cost between firms that borrow from an old financial institution and firms that use a new institution is \$5,650 (column 1). The difference in means is significant at the 1% level with heteroskedasticity-robust standard errors. The mean of *Cost* is \$3,160, and the standard deviation of *Cost* is \$9,897. Consequently, the \$5,650 difference in mean *Cost* equals 0.57 of the standard deviation of *Cost* and 1.79 times the mean. With the addition of controls (columns 2 to 4) the R^2 rises to 0.25, yet the coefficient on *NoRelationship* does not change much. The coefficient remains positive and significant at α=1%.

The regressions in Table 3 control for the maximum length of all the financial relationships maintained by a firm with financial institutions. The significance of the coefficient on *NoRelationship* in the presence of this control suggests that the effect of *NoRelationship* is not driven by poorly performing firms' inability to maintain a long relationship. The positive coefficients on *Amount of Credit Granted* and *Maturity* suggest that larger, more complex loans require higher origination costs. Similarly, the positive coefficients on *Mortgage for Business Purposes* and *Vehicle Loan for Business Purpose*s indicators suggest that loans based on sizable collateral that requires appraisal also necessitate higher up-front costs. The positive coefficients on *Interest Rate* and the negative coefficients on the *D&B Credit Score* indicators suggest that riskier firms are

[15] See footnote 1 on page 1.

asked to pay higher origination costs. This finding could be driven by lenders' use of origination costs to price risk. This finding may also be attributable to riskier firms' need to provide a lender with more complex financial documentation and a larger number of collateral appraisals in order to secure a loan.

4.2 Causal Inference

This paper aims to assess the causal effect of borrowing from a new lender on origination costs. But, what does *causality* mean in the context of the present analysis? In general, a treatment is said to have a causal effect on an outcome if a change in the treatment result coincides with in a change in the outcome when all other factors that influence the outcome are held constant. Following the potential outcomes framework of the Rubin model, a treatment is said to have a causal effect on an outcome if (1) the outcome is observed to differ between treated and untreated subjects and (2) treatment is assigned such that the potential (expected) outcome given treatment is the same for the subjects in the treated group and the subjects in the untreated group.[16] Also, the potential outcome given no treatment should be the same for the treated group and the untreated group. The treatment is observed to have an effect while the treatment and control groups are identical in how they should respond to the presence and the absence of treatment. Causal effects are most commonly estimated using a laboratory experiment wherein subjects are randomly assigned to be in the treated group or the untreated group (the control group). Random assignment ensures that subjects in the treated group and subjects in the untreated group are statistically identical.

With respect to the present analysis of small business lending, one could infer that borrowing from a previously unused financial institution has a causal effect on loan origination costs if one were able to run the following experiment. Obtain a list of firms surveyed in the 2003 SSBF that have at least one existing relationship with a lending institution. Randomly assign some of those firms to apply for a loan from their old financial institution, and assign the remainder of the firms to apply for a loan from a new institution. Then, observe the firms' loan origination costs. If the firms that borrow from a new lender have higher loan origination costs than firms that borrow from their old lender, then we can infer that borrowing from a new lender has a positive causal effect. When I state that *NoRelationship* should have a causal effect equal to a number b, I mean to say that if the above experiment were conducted we would observe that, on average, among firms for which *NoRelationship*=1 *Cost* is b greater than among firms for which *NoRelationship*=0.

We can think of causality in aother way. Suppose we again obtained a list of firms surveyed in the 2003 SSBF that have at least one existing relationship with a lending institution. Suppose we asked each firm to apply to one of its previously used lenders and then we also asked each firm to apply to a new lender (randomly selected by the experimenter). In this experiment, the treatment group doubles as the control group, so the distributions of firm characteristics are identical in both groups. Only the lenders differ. This experiment should also allow us to identify the causal effect of borrowing from a new financial institution. As before, when I claim that *NoRelationship* has a

[16] See Angrist and Pischke (2009) for a comprehensive discussion. Also, see Rubin (1974, 1977).

causal effect equal to a number b, I mean that if the second experiment were conducted, we would observe that *Cost* among loan offers from firms' old institutions would be b greater than *Cost* among loan offers from new institutions.

More rigorously, consider the following population model for loan origination costs:

$$Cost_i = \kappa + \beta_0 NoRelationship_i + \mathrm{W}_i'\Omega + \nu_i \tag{2}$$

Here, W is the set of all variables that affect *Cost* and are correlated with *NoRelationship*, and ν is an index (linear combination) of all other variables that are uncorrelated with *NoRelationship* and W. $\mathbf{E}[\nu|\mathrm{W}, NoRelationship]=0$. The coefficient β_0 is the causal effect of borrowing from a previously unused lender. The model explicitly assumes that the causal effect of *NoRelationship* is identical for all firms. If all the variables in W were observed by the econometrician, then β_0 could be estimated using simple OLS.

4.3 Potential Sources of Endogenous Selection

The results presented in Table 1 suggest that borrowing from a new financial institution causes firms to pay higher loan origination costs. However, since the estimation relies on observational data where treatment — borrowing from a new lender — is not randomly assigned, the results from Table 1 can only suggest causality. Among small business borrowers selection into borrowing from a new lender may be correlated with unobservable variables that simultaneously affect loan origination costs. The presence of unobservables that determine *Cost* and are correlated to *NoRelationship* would confound my estimation of causal effects.

The regression coefficient on *NoRelationship* could be biased due to endogenous selection on firms' ability to borrow from a new lender. High credit quality firms may be more likely to switch lenders while simultaneously taking on bigger or more complex loans that require larger origination costs. Such selection would bias the estimated coefficient upward, suggesting a positive causal relationship where none exists. On the other hand, high-credit-quality firms could be more likely to turn to a new lender, and they may be asked to pay fewer origination fees. Higher quality firms may have less need to provide detailed documentation, or they may simply pay fewer fees because banks use origination cost to price risk. This would lead to a negatively biased coefficient that underrepresents the true causal effect.

My estimation could also be confounded by endogenous selection on firms' ability to maintain a relationship with a financial institution. Low-credit-quality firms may be less able to maintain a relationship with a financial institution, and they may be more likely to apply to a new lender for a loan. If such firms have higher loan origination costs due to their risk or their need to produce extensive financial documentation, the coefficient on *NoRelationship* would be biased upward. On the other hand, if such firms take out smaller, less complex loans that require fewer up-front fees, then the coefficient would be biased downward.

My estimation could also be biased due to selection on loan contract terms. It could be that small business borrowers switch to a new financial institution if the new lender offers them better terms such as a larger loan size, longer maturity, lower interest rate, or less collateral. If better loan terms require more due diligence and higher origination

fees, then my estimated coefficient would be biased upward, as it would reflect selection on loan terms and not the causal effect of borrowing from a new lender. It is also possible that firms are more likely to turn to a new lender when the origination costs offered by the new lender are lower. Some firms might choose a new lender because the new lender offers a loan with a low origination cost. In this case my coefficient would be biased downward, severely underrepresenting the causal effect of borrowing from a previously unused bank.

One could suppose that ceteris paribus small businesses prefer to borrow form a previously unused financial institution. If that were the case, then firms could be selecting to borrow from a previously used financial institution if and when the old institution offers the firm low loan origination costs. In this scenario, my estimation would be confounded by something similar to reverse causation, and consequently the estimated coefficient $\hat{\beta}$ would be biased upward. Fortunately, this scenario is highly unlikely. The SSBF asks firms to list reasons why they chose to apply to the specific lender who provided them with their most recent loan. Fifty-seven percent (57%) of all respondents who mentioned reasons cited having a prior relationship with the institution as one of the reasons behind their choice to apply to the lender. Having a prior relationship with a financial institution was cited as the number one reason for selecting a lender. In contrast, only 1.2% of respondents cited a wish to borrow from a new lender as a reason for why they originally applied to the provider of their most recent loan.

It is impossible that firms would borrow from a new lender in response to the lender offering them a high origination cost. It seems highly unlikely that borrowing from an old financial institution is driven by firms' lopsided sensitivity to the origination costs offered by their old institutions. It seems much more likely that firms are more sensitive to the loan terms offered by new lenders than to the terms offered by their old financial institutions. If firms select lenders based on expected origination costs, then firms are probably more likely to borrow from a new lender if the new lender offers an especially low origination cost. This would imply that the estimated $\hat{\beta}$ is biased downward, understating the true causal effect of borrowing from a new financial institution.

5. A More Comprehensive Regression Analysis

5.1 Assessing the Possibility of a Causal Effect In the Presence of Many Controls

I extend the analysis in Table 3 by adding many more controls to my estimation of Equation 1. I use the richness of the SSBF data to control for many important firm characteristics and for the loan terms of borrowers' most recent loan. The results of my estimation of Equation 1 with the added controls are presented in Table 4. Table 4 also details all the control variables. The regressions presented in Table 4 control for firms' financial relationships. They also control for firm age, size, financial state, credit quality, and organization. The regressions also control for firms' local geographic markets, local banking market concentration, and macroeconomic variables. I also control for the characteristics of firms' most recent loans. I also include a set of dummy variables that

indicate the reasons cited by firms for applying to the lender that granted their most recent loan. Lastly, I include a set of 2-digit SIC code industry fixed effects.

The addition of controls raises the R^2 to 0.41 but does not diminish the size or the magnitude of $\hat{\beta}$. Indeed, the estimated $\hat{\beta}$ rises with the addition of controls and always remains significant at the 1 percent level. My controls are powerful predictors of firm finances and borrowing behavior, yet in their presence the coefficient remains remarkably stable. This suggest that firms' selection to borrow from a previously unused lender is largely independent from firms' most salient characteristics.

If the simple single-variable analysis in column 1 suffers from omitted variable bias, the bias seems to be working against the coefficient, not inflating it. This suggest that, on average, firms' selection to borrow from a new financial institution is negatively correlated with unobservables that increase loan origination costs. If the unobserved factors determining *Cost* are largely similar to the observables, then the lowest estimate of β, $\hat{\beta}$ =\$5,650, is likely below the true causal effect of borrowing from a new financial institution.

5.2 A Sample With Less Potential for Endogenous Selection

As a robustness check, I reestimate the models described above using a subsample of firms that is less likely to suffer from endogenous selection. I select firms that are (1) less likely to have difficulties borrowing from a new lender and (2) less likely to have difficulties maintaining a financial relationship. I select firms that have one or more relationships with a financial institution that is at least three years long. I also exclude those firms that state that they applied to the lender of their most recent loan because they had difficulty obtaining credit from other institutions. I exclude firms that had a loan application denied during their most recent attempt to procure a loan. I include only firms that have a D&B credit score above 50 (out of 100). Lastly, I include only those firms that had two or more financial relationships at the time of their most-recent loan application.

The coefficients from the estimation of Equation 1 on this subsample are presented in Table 5. The coefficients in Table 5 are all larger than the baseline estimate of 5.65 and are all significant at the 1% level. The univariate estimate of the difference in means between firms that borrow from a previously used institution and firms that borrow from a new institution is equal to \$7,150. The coefficient declines with the addition of controls to a low of \$6,140. The analysis in Table 5 strongly suggests that our primary finding is not driven by select firms' inability to borrow from a new lender, and it is not driven by some firms' difficulty with maintaining a financial relationship.

6. A Natural Quasi-Experiment — Firms That Select a Lender Based on Proximity

The analysis presented thus far suggests that the β coefficient estimated using OLS should not be too different from the true causal effect β_0. But, we cannot rule out with certainty the influence of endogenous selection. To obtain a more precise estimate of the causal effect of borrowing from a new lender, I look for a natural quasi-experiment wherein selection to borrow from an old lender is close to randomly assigned. Among the reasons cited by firms as to why they chose to apply to the lender of their most recent loan is the lender's proximity. Forty-seven (47) firms state that they chose to borrow from the lender of their most recent loan because of the lender's proximity to their location. A small business has no control over the opening or closing of financial institutions and their branches. For a small business that chooses a lender based on proximity, lender selection is driven by financial institutions' unpredictable entry and exit into the firm's local market. Among these firms, lender selection is not driven by firm characteristics nor is it driven by lender characteristics other than proximity.

The 2003 SSBF survey questionnaire asks each surveyed firm to explain why the firm chose the lender that provided the firm with its most-recent loan. The survey asks each firm to explicitly state the top three reasons that influenced the firm's decision to apply for credit from the issuer of the firm's most recent loan.[17] The survey questionnaire directs the questioner to explicitly ask the survey taker "What factors influenced the firm's decision to apply for credit from [name of provider of most-recent loan]?" The questioner is then directed to "record verbatim response." It may be possible that firms who self-identify as distance-oriented consider other factors in addition to lender proximity. But, because the survey takers volunteer their responses without their responses being framed by a constrained set of multiple-choice answers, it is almost certain that respondents' answers are accurate.

I select those firms that cite proximity as a reason for selecting a lender and estimate Equation 1 without controls on this subsample of the data. There are only 47 such firms in my data. So, to allow for a rich set of controls, I also estimate a regression model that combines data on these 47 firms with data on other firms. This regression is as follows.

$$Cost_i = \kappa + \alpha NotDist_i \cdot NoRelationship_i + \pi Dist_i + \theta Dist_i \cdot NoRelationship_i + X_i'\gamma + \varepsilon_i \qquad (3)$$

Here $Dist_i$ is a dummy variable indicating if a firm cites distance to lender as a reason for selecting the lender. $NotDist_i = 1 - Dist_i$. The terms $NotDist_i \cdot NoRelationship_i$ and $Dist_i \cdot NoRelationship_i$ are interaction terms. Alpha (α), π, and θ are coefficients. While Model 3 looks similar to a difference-in-differences model, it is not a difference-in-differences model. The coefficient μ is not a difference in differences interaction coefficient. Model 3 simply combines Model 2 for firms having $Dist=0$ and Model 2 for firms having $Dist=1$ into a single equation. Without controls, the coefficient θ is numerically identical to the coefficient β in Equation 2 when Equation 2 is estimated only using firms that have $Dist=1$.

[17] See question for variable "MRL9" on pages 158-159 of the *2003 Survey of Small Business Finances Survey Questionnaire* (www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html).

Among firms that cite lender proximity as a reason for selecting a lender, a firm's selection to borrow from a new lender should be less endogenous than among the general population of small businesses. Yet, firms that cite proximity as a reason for choosing a lender may also consider other reasons. The SSBF does allow a respondent to cite up to three different reasons for their choice. I strengthen the precision of my analysis by focusing on firms that simultaneously (1) cite distance to lender as a reason for lender selection, (2) are less likely to have difficulties borrowing from a new lender, and (3) are less likely to have difficulties maintaining a financial relationship. Following the same criteria as laid out in Section 5.2, I identify 23 firms that cite distance as a reason for selecting a lender and are less constrained in their ability to select a lender of their choice. I estimate Model 2 for this small set of firms. To allow for a rich set of controls, I also estimate a model almost identical to Model 3 (above). The model is presented below in Equation 4.

$$Cost_i = \kappa + \tilde{\alpha} NotDLC_i \cdot NoRelationship_i + \tilde{\pi} DLC_i + \tilde{\theta} DLC_i \cdot NoRelationship_i + X_i'\gamma + \varepsilon_i \qquad (4)$$

Here DLC_i is a dummy variable indicating if a firm both cites distance to lender as a reason for selecting the lender and is less constrained. $NotDLC_i = 1 - DLC_i$. The terms $NotDLC_i \cdot NoRelationship_i$ and $DLC_i \cdot NoRelationship_i$ are interaction terms. Alpha with a tilde ($\tilde{\alpha}$), $\tilde{\pi}$, and $\tilde{\theta}$ are coefficients.

The results from the estimations just described are presented in Table 6. Column 1 shows that within the small group of 47 firms that select a lender based on proximity *NoRelationship* by itself explains 42 percent of all variation in *Cost* — R^2=0.42. The coefficient on *NoRelationship* is equal to 13.32 ($13,320) and is significant at the 5% level with robust standard errors. With homoscedastic standard errors, the coefficient has a *t*-statistic of 5.75 and a *p*-value of less than 0.000001. Column 3 shows that the addition of controls slightly reduces the coefficient to 10.64 ($10,640). The coefficient remains significant at 5%. Borrowing from a previously unused financial institution is likely to have at least some small effect on loan contract terms, and loan contract terms are also correlated to origination costs. Consequently, it is understandable that controlling for loan contract terms may slightly attenuate $\hat{\mu}$.

When I restrict my focus to firms that both consider lender proximity and are less constrained in their lender selection, my estimated results are almost identical to the results described in the preceding paragraph. The estimates in columns 4, 5, and 6 of Table 6 are almost exactly the same as the estimates in columns 1, 2, and 3. This strongly suggests that within the set of firms that select a lender based on distance, borrowing from a new financial institution is uncorrelated to those covariates that are not themselves influenced by the act of borrowing from a new lender. Among firms that care about the proximity of a lender, selection to borrow from a previously used financial institution does indeed seem strongly uncorrelated with firm characteristics.

The results presented above strongly suggest that borrowing from a previously-unused financial institution has a causal effect on loan origination costs among those firms that select a lender based on distance. Unfortunately such firms make up a very small slice of the population of SSBF-surveyed small businesses. Firms that select a lender based on proximity make up only 3.73% of the 1,261 firms in my data. So, one cannot be certain that the existence of a causal effect of borrowing from a new lender

seen among these distance-oriented firms necessarily implies the existence of a causal effect for the average firm sampled in the SSBF.

Indeed, it is possible that the kinds of firms that cite lender proximity are located in rural or suburban areas where there are few banks and where the distance between neighboring banks may be many miles. It is also possible that the kinds of firms that can focus on distance to lender are likely to frequently use many different services offered by a lender. They may also be the kinds of firms that require frequent and perhaps personal interaction with their lender. They may need to take out complicated loans that need to be frequently renewed or renegotiated. Such loans may need to be linked to the firms' checking and savings accounts. I leave more detailed analysis of what kinds of firms borrow from the closest lender to later research.

While the quasi-experiment just described is informative, the small number of firms it analyzes may not be representative of the general population of SSBF small businesses. To estimate the causal effect of borrowing from a new financial institution among the general population of firms sampled in the SSBF, I need to make use of another econometric analysis. I turn to this next.

7. Estimation of Causal Effects Using the Method of Altonji, Elder, and Taber.

While the coefficients described in Section 5 and Table 2 are illuminating, accepting them as estimates of causal effects would require me to assume that borrowing from a new lender is strictly exogenous. The estimations of Section 5 contain an abundance of controls, yet no set of controls is ever perfect. A firm's selection to borrow from a new financial institution may be related to unobserved factors that simultaneously affect the firm's origination costs. To correct for the presence of endogenous selection in my econometric model I implement the recently-developed method of Altonji, Elder, and Taber (2002, 2005). The method of Altonji, Elder, and Taber (AET) allows me to leverage the large number of covariates at my disposal to identify a lower bound on the causal effect in question.

7.1 Model Description

In their 2005 article (Journal of Political Economy) Altonji, Elder, and Taber model a binary outcome determined by a possibly endogenous binary treatment effect. The authors consider a selection equation for the endogenous binary treatment effect that is jointly determined with the outcome equation. The authors estimate a two-equation bivariate (two-equation) maximum likelihood model that maximizes the joint likelihood of two probit likelihoods functions. Altonji et al. obtain a point estimate of the treatment effect by specifying a restriction on the correlation between the residuals in the selection equation and the outcome equation.

I adapt Altonji, Elder, and Taber's approach to allow for a continuous outcome variable. I specify a simple bivariate model as follows.

$$Cost_i = \kappa + \beta NoRelationship_i + X_i'\gamma + \varepsilon_i \tag{5.1}$$

$$NoRelationship_i = \mathbf{1}\{\lambda + X_i'\psi + u_i > 0\} \tag{5.2}$$

$$\begin{bmatrix}\varepsilon_i \\ u_i\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2 & \rho\sigma \\ \rho\sigma & 1\end{bmatrix}\right) \tag{5.3}$$

Here, *Cost* is the outcome of interest, *NoRelationship* is the potentially endogenous treatment variable, and β is the coefficient of interest. X is a vector of controls, identical in both the outcome and the treatment equation, and γ and ψ are a vectors of coefficients. Kappa (κ) and λ are constant terms, and ε and u are error terms. $\mathbf{1}\{\,\}$ denotes a binary indicator function, and $\mathcal{N}$ indicates a bivariate normal distribution with mean zero. The variance of ε is σ^2; the variance u is 1; and the correlation between ε and u is equal to ρ.

Although origination cost, like all positive variables, is truncated from below at zero, I use a Normal (Gaussian) distribution for the outcome likelihood function. The coefficients calculated by ML estimation of a linear model with Normal errors are numerically identical to the coefficients estimated by OLS regression. Consequently, like OLS, linear Normal ML consistently estimates the conditional expectation of the outcome variable — the population regression function. Like OLS, linear Normal ML is consistent even when the error term is heteroskedastic and not Normally distributed. Like OLS, linear Normal ML is very robust.

In contrast, the Tobit model imposes very strict assumptions on the data generating process of input data. The Tobit model requires that error terms be strictly i.i.d Normal; the Tobit model is not robust to heteroskedasticity. More importantly, the Tobit model imposes exceedingly strict assumptions on the data generating process of the outcome variable. The Tobit model decomposes the data generating process into (1) a process that determines if an observation lies on the truncation boundary of 0 and (2) a process that determines an observation's value if it lies above the truncation boundary. Then, the Tobit model imposes the strict assumption that the same statistical process, identified by one specific functional form and one specific Normal distribution, describes both (1) whether an observation lies on the truncation boundary and (2) the observed value of an observation if it lies above the boundary.

From another perspective, the error term of a long, multi-variable linear model is very likely to be Normally distributed even if the outcome being modeled is not itself Normally distributed. Consider the equations below that describe an outcome y decomposed into a linear combination of K components — a constant plus $K-1$ covariates.

$$y = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 + \cdots + \delta_{K-1} x_{K-1} + e \tag{6.1}$$

$$\begin{aligned} e_i &= y - (\delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 + \cdots + \delta_{K-1} x_{K-1}) \\ &= y - K\cdot\widehat{\mathbf{E}}[\delta_k x_k] = \widehat{\mathbf{E}}\left[y - \overline{\delta}_k x_k\right] \end{aligned} \tag{6.2}$$

Here, $x_0{=}1$, $\delta_0 x_0{=}\delta_0$, and $\delta_k = \overline{\delta}_k / K$. Equation 6.1 shows a long, multi-variable linear model, and Equation 6.1 rearranges the model to define the error term. Equation 6.2 reveals that the error term is the arithmetic mean of K random variables $y - \overline{\delta}_k x_k$. (The symbol $\widehat{\mathbf{E}}[\cdot]$ represents an arithmetic mean.) The Central Limit Theorem states that e

should be Normally distributed even if the random variables $y - \overline{\delta}_k x_k$ are not identically distributed and not completely independent.

In my estimation, the vector X contains all the variables listed in Table 2 with the following exceptions. A few variables had to be omitted from my vector of controls because their inclusion created convergence problems in estimating the AET model. These excluded variables are: *Financial Statements Audited*, *Sales Radius Indicators*, *Mean Distance to Financial Institutions*, and *Banking Market HHI Indicators*.

7.2 Identification in the Altonji, Elder, and Taber Model

Unlike traditional approaches to selection problems, which typically rely on one or more exclusion restrictions to achieve identification, the methodology of Altonji, Elder, and Taber does not require exclusion restrictions. Instead, Altonji et al. achieve identification through an additional restriction on ρ, the correlation between ε and u. The authors' approach, in the context of Model 5, is as follows. As before, for ease of exposition, some subscripts are omitted. The binary treatment variable *NoRelationship* is determined by an underlying, latent variable *NoRelationship** in the usual way:

$$NoRelationship_i = \mathbf{1}\{NoRelationship^* > 0\} \tag{7}$$

Using linear projection, project *NoRelationship** onto the linear combination of controls in the outcome equation, $X'\gamma$, and the error term in the outcome equation, ε, as follows.

$$\text{Proj}\left(NoRelationship^* \middle| X'\gamma, \varepsilon\right) = \phi_0 + \phi_{X'\gamma} X'\gamma + \phi_\varepsilon \varepsilon \tag{8}$$

Altonji, Elder, and Taber argue that if the variables included in X are chosen at random, then the projection coefficients $\phi_{X'\gamma}$ and ϕ_ε should be equal.[18] The intuition behind the authors' argument is as follows. Consider the set of all variables that determine *Cost*, and allow this set of underlying variables to be sufficiently large. Suppose this set were randomly partitioned into two distinct subsets of variables, $\tilde{X}_a$ and $\tilde{X}_b$, where

$$Cost = \kappa + \beta NoRelationship + \tilde{X}_a'\tilde{\gamma}_a + \tilde{X}_b'\tilde{\gamma}_b \tag{9}$$

Then, in the linear projection of any variable onto $\tilde{X}_a'\tilde{\gamma}_a$ and $\tilde{X}_b'\tilde{\gamma}_b$ the projection coefficients on $\tilde{X}_a'\tilde{\gamma}_a$ and $\tilde{X}_b'\tilde{\gamma}_b$ should be approximately equal. As the number of variables in the set of all underlying variables that determine *Cost* goes to infinity, the difference between $\tilde{X}_a'\tilde{\gamma}_a$ and $\tilde{X}_b'\tilde{\gamma}_b$ converges to zero. In Expression 8, $X'\gamma$ is analogous to $\tilde{X}_a'\tilde{\gamma}_a$, and ε is analogous to $\tilde{X}_b'\tilde{\gamma}_b$. Consequently, if X contains a large number of variables, $\phi_\varepsilon \cong \phi_{X'\gamma}$.

However, as Altonji, Elder, and Taber state, the assumption that the elements of X are selected at random is "not to be taken literally." Like Altonji, Elder, and Taber, I have selected the control variables in X "with an eye towards reducing bias in single

[18] The restriction $\phi_\varepsilon = \phi_{X'\gamma}$ can be compared to the standard OLS assumption of $\phi_\varepsilon = 0$.

equation estimates." Like Altoji et al., I have attempted to select variables that should be "better than random" in their ability to explain *Cost* and in their ability to explain *NoRelationship*. Assuming, without loss of generality, that $\phi_\varepsilon \geq 0$, it seems more likely that $0 \leq \phi_\varepsilon \leq \phi_{X'\gamma}$. More generally:

$$0 \leq \left|\phi_\varepsilon\right| \leq \left|\phi_{X'\gamma}\right| \tag{10}$$

This inequality translates into a very useful condition that allows me to place bounds on the possible values of ρ. The condition is:

$$0 \leq |\rho| \leq \sigma \left| \frac{\operatorname{cov}\left(X'\psi, X'\gamma\right)}{\operatorname{var}\left(X'\gamma\right)} \right| \tag{11}$$

Note that if the variables in X are selected to effectively control for any endoge-nous relationship between *NoRelationship* and *Cost*, then endogenous selection is not a concern and ρ=0. At the other extreme, if X is truly randomly selected, then

$$\rho = \sigma \left(\frac{\operatorname{cov}\left(X'\psi, X'\gamma\right)}{\operatorname{var}\left(X'\gamma\right)} \right) \tag{12}$$

and Model 5 can be properly estimated by imposing Condition 12. I use the method of maximum likelihood to estimate Model 12. First, I estimate the model under the assumption that ρ=0. This yields the same β coefficient as OLS. Then, I estimate the model under the explicit imposition of Condition 12, obtain β_{AET}, and compare the two estimates. Just as the true value of ρ should be between 0 and $\rho = \sigma \cdot \operatorname{cov}\left(X'\psi, X'\gamma\right) \big/ \operatorname{var}\left(X'\gamma\right)$, the true value of the causal effect being estimated should lie between β_{OLS} and β_{AET}.

7.3 Estimation Results

I estimate the Altonji, Elder, and Taber model using two different specifications. First, I estimate the model using only the variables in Column 3 of Table 4 as covariates. Second, I estimate the model using my largest set of control variables — all the variables in Column 4 of Table 4. Table 7 displays my results. The top panel of Table 7 shows that a firm's use of a single lender is only weakly correlated with the residual of the outcome equation; $-0.074 \leq \hat{\rho} \leq -0.039$. Since, the values of $\hat{\rho}$ displayed in columns 3 and 4 are estimated under the relatively strong assumption of random selection of observables, it is reasonable to conclude that ρ most likely lies between -0.074 and 0 ($-0.074 \leq \rho \leq 0$). Because $\hat{\rho}$ is negative, β_{OLS} is revealed to be the lower bound of the causal effect we are aiming to estimate. Thus, the causal effect of borrowing from a financial institution is estimated to be no less than \$5,740.

8. The Implausibility of Loyalty Discounts

The analysis presented thus far shows that borrowing from a new financial institution causes a firm to pay higher loan origination costs than borrowing from an institution with which a firm has a preexisting relationship. But, is the documented relationship effect the result of the information content of lender-borrower relationships? Or, is the relationship effect nothing more than loyalty-based discounting? In this section I argue that it is the former. Lender-borrower relationships give banks in-depth knowledge of their borrowers, and that knowledge reduces the need for an extensive and expensive loan origination process.

To obtain empirical support this claim, I look at firms that cite low prices as a reason for their selection to apply to the bank that issued them their most-recent loan. Firms that cite low prices as a reason for their choice of lender state that they selected their lender because the lender offered low loan interest rates or low fees (or both). I separately estimate the following two regressions:

$$InterestRate_i = \kappa + \alpha Price_i + \beta NoRelationship_i + X_i'\gamma + \varepsilon_i \tag{13}$$

$$Cost_i = \kappa + \alpha Price_i + \beta NoRelationship_i + X_i'\gamma + \varepsilon_i \tag{14}$$

where *InterestRate* is the interest rate charged the firm on its most recent loan, *Price* is a dummy variable indicating if a firm cites prices as a reason for selecting its lender, and the remaining terms are as described before. The estimations of equations 13 and 14 are presented in Tables 8 and 9, respectively. Tables 8 and 9 reveal that price-conscious firms pay significantly lower loan interest rates than non-price-conscious firms. In contrast, being price-conscious is not correlated to loan origination costs. Moreover, Table 8 reveals that borrowing from an old lender does not result in lower interest rates.

These findings could be the result of banks' preference to respond to demand-side pressures by discounting interest rates instead of origination costs. Such a scenario would contradict the claim that banks offer loyalty discounts on origination costs while not offering loyalty discounts on interest rates. I do not believe that banks are likely to prefer to discount interest rates over origination fees. I believe banks should be indifferent between all possible discount programs and should offer discounts based on customer preferences.

It is more likely that the results in Tables 8 and 9 are driven by customer preferences. Tables 8 and 9 strongly suggest that price-consciousness is manifested as a sensitivity to loan interest rates. Price-conscious firms are more sensitive to loan interest rates than to loan origination fees. If this relative sensitivity, privileging interest rates to origination costs, is representative of the majority of small business borrowers, then offering a discount on interest rates is much more likely to affect demand than offering a discount on origination costs. Yet, as Table 8 shows, borrowing from an old lender does not result in a lower interest rate. Offering loyalty discounts on origination costs while not offering discounts on loan interest rates seems inconsistent with any pricing strategy meant to affect customer demand.

9. Conclusion

This paper studies the relationship between the borrowing choices made by small businesses and their loan origination costs. I attempt to assess the causal effect of borrowing from a new lender on small businesses' up-front closing costs. Simple cross sectional regression shows that small businesses that turn to a new financial institution pay $5,650 to $6,740 more in closing costs than firms that return to a previously-used institution.

Next, I study a natural quasi-experiment wherein the treatment of borrowing from a new financial institution is close to randomly assigned. I consider a unique group of small businesses that select a lender based on the lender's distance to the firm. Data on firms that select a lender based on proximity reveal that borrowing from a new financial institution significantly raises a firm's loan origination costs.

Lastly, I estimate the causal effect of borrowing from a new financial institution among the general population of firms. I use the method of Altonji, Elder, and Taber (2002, 2005) to identify a lower bound on the size of the causal effect of borrowing from a new financial institution. The method of Altonji, Elder, and Taber reveals that even after accounting for endogenous selection, the causal effect of borrowing from a new lender remains positive and significant.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2005, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113(1), 151–184.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2002, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," Manuscript (April), Department of Economics, Northwestern University.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2008, "Using Selection on Observed Variables to Assess Bias from Unobservables When Evaluating Swan-Ganz Catheterization," *American Economic Review*, 98(2), *Papers and Proceedings of the One Hundred Twentieth Annual Meeting of the American Economic Association*, 345–350.

Agarwal, Sumit and Faye H. Wang, 2009, "Perverse Incentives at the Banks? Evidence from a Natural Experiment," Working Paper, Federal Reserve Bank of Chicago.

Angrist, Joshua D. and Jörn-Steffen Pischke, 2009, Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.

Berg, Tobias, Anthony Saunders, and Sascha Steffen, 2013, "The Total Costs of Corporate Borrowing in the Loan Market: Don't Ignore the Fees," Working Paper.

Berg, Tobias, Anthony Saunders, and Sascha Steffen, 2015, "The Total Costs of Corporate Borrowing in the Loan Market: Don't Ignore the Fees," Working Paper (forthcoming in the *Journal of Finance*).

Berg, Tobias, Manju Puri, and Jörg Rocholl, 2014, "Loan Officer Incentives, Internal Ratings and Default Rates," Working Paper.

Berger, Allen N. and Gregory F. Udell, 1995, "Relationship Lending and Lines of Credit in Small Firm Finance," *Journal of Business* 68, 351–381.

Bharath, Sreedhar T., Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan, 2011, "Lending Relationships and Loan Contract Terms," *Review of Financial Studies* 24(4), 1141–1203.

Bodenhorn, Howard, 2003, "Short-Term Loans and Long-Term Relationships: Relationship Lending in Early America," *Journal of Money, Credit, and Banking* 35(4), 485–505.

Boot, Arnaud W.A., 2000, "Relationship Banking: What Do We Know?," *Journal of Financial Intermediation* 9, 7–25.

Brick, Ivan E. and Darius Palia, 2007, "Evidence of Jointness in the Terms of Relationship Lending," *Journal of Financial Intermediation* 16, 452–476.

Cole, Rebel A., 1998, "The Importance of Relationships to the Availability of Credit," *Journal of Banking & Finance* 22, 959–977.

DeGryse, Hans and Patrick Van Cayseele, 2000, "Relationship Lending Within a Bank-Based System: Evidence From European Small Business Data, *J. of Fin. Intermediation* 9, 90–109.

DeGryse, Hans, Moshe Kim, and Steven Ongena, 2009, *Microeconometrics of Banking*, Oxford University Press.

Elsas, Ralf and Jan Pieter Krahnen, 1998, "Is Relationship Lending Special? Evidence From Credit-File Data in Germany," *Journal of Banking & Finance* 22, 1283–1316.

Freixas, Xavier and Jean-Charles Rochet, 2008, *Microeconomics of Banking, 2nd Ed.*, MIT Press.

Hellmann, Thomas, Laura Lindsey, and Manju Puri, 2008, "Building Relationships Early: Banks in Venture Capital," *Review of Financial Studies* 21(2), 513–541.

Ioannidou, Vasso and Steven Ongena, 2010, "'Time for a Change': Loan Conditions and Bank Behavior When Firms Switch Banks," *Journal of Finance*, 65(5), 1847–1877.

Iyer, Rajkamal and José-Luis Peydro, 2011, "Interbank Contagion at Work: Evidence from a Natural Experiment," *Review of Financial Studies* 24(4).

Iyer, Rajkamal and Manju Puri, 2012, Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks," *American Economic Review* 102(4), 1414–1445.

Iyer, Rajkamal, Manju Puri, and Nicholas Ryan, 2015, "A Tale of Two Runs: Depositor Responses to Bank Solvency Risk," Working Paper.

Jiangli, Wenying, Haluk Unal, and Chiwon Yom, 2008, "Relationship Lending, Accounting Disclosure, and Credit Availability During the Asian Financial Crisis," *Journal of Money, Credit and Banking* 40(1), 25–55.

Klemperer, Paul, 1987, "Markets With Consumer Switching Costs," *Quarterly Journal of Economics*, May.

Krishnan, Karthik, Debarshi K. Nandy, and Manju Puri, 2014, "Does Financing Spur Small Business Productivity? Evidence from a Natural Experiment," *Review of Financial Studies*, RFS Advance Access, December 30.

Ongena, Steven and David C. Smith, 2000, "Bank Relationships: A Review," in *Performance of Financial Institutions: Efficiency, Innovation, Regulation*, P. Harker and S. Zenios (eds.), Cambridge University Press.

Oster, Emily, 2014, "Unobservable Selection and Coefficient Stability: Theory and Evidence," Working Paper.

Petersen, Mitchell A., 1999, "Banks and the Role of Lending Relationships: Evidence From the U.S. Experience," Working Paper, Kellogg Graduate School of Management, Northwestern University.

Petersen, Mitchell A., and Raghuram G. Rajan, 1994, "The Benefits of Lending Relationships: Evidence From Small Business Data," *Journal of Finance* 49, 3–37.

Puri, Manju, Jörg Rocholl, and Sascha Steffen, 2011, "On The Importance of Prior Relationships in Bank Loans to Retail Customers," Working Paper No. 1395, European Central Bank.

Puri, Manju, Jörg Rocholl, and Sascha Steffen, 2011, "Global Retail Lending in the Aftermath of the US financial Crisis: Distinguishing Between Supply and Demand Effects," *Journal of Financial Economics* 100, 556–578.

Puri, Manju, Jörg Rocholl, and Sascha Steffen, 2013, "What Kinds of Bank-Client Relationships Matter in Reducing Loan Defaults and Why?," Working Paper.

Rubin, Donald B., 1974, "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology* 66(5), 688–701.

Rubin, Donald B., 1977, "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational Statistics* 2(1), 1–26.

Sharp, Byron, and Anne Sharp, "Loyalty Programs and Their Impact on Repeat-Purchase Loyalty Patterns," *International Journal of Research in Marketing* 14, 473–486.

"2003 Survey of Small Business Finances Technical Codebook," 2008, *Survey of Small Business Finances Group*, Division of Research and Statistics, Board of Governors of the Federal Reserve System, https://www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html.

"2003 Survey of Small Business Finances Annotated Survey Questionnaire With Forward and Appendices," 2008, *Survey of Small Business Finances Group*, Division of Research and Sta-

tistics, Board of Governors of the Federal Reserve System, https://www.federalreserve.gov/pubs/oss/oss3/ssbf03/ssbf03home.html.

Udell, Gregory F., 2008, "What's In a Relationship? The Case of Commercial Lending," *Business Horizons* 51, 93–103.

Uzzi, Brian, 1999, "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit firms Seeking Financing," *American Sociological Review* (64), 481–505.

Ziebarth, Nicolas L., 2013, "Identifying the Effects of Bank Failures From a Natural Experiment in Mississippi During the Great Depression," *American Economic Journal: Macroeconomics*, 5(1), 81–101.

Table 1. Definitions and Summary Statistics. Definitions and statistics for variables used in estimation. Data consist of 1,261 observations on 1,261 loans. Some observations are incomplete. Except for macroeconomic interest rates, all data come from the 2003 Survey of Small Business Finances.

Variable Name	Variable Definition	Mean	St.Dev.	Min	Max
Dependent Variable					
Origination Cost of Most-Recent Loan ($thousands)	Total dollar amount of fees paid to apply for and obtain loan.	3.16	9.90	0	66.18
Treatment Variable					
No Relationship (New Lender)	Binary indicator = 1 if firm has no prior relationship with lender at loan origination.	0.067	0.25	0	1
Financial Relationships					
Max Relationship Length (years)	Maximum length of a relationship with a financial institution.	15.46	18.38	0	96.25
Number of Loans	Number of loans obtained by the firm from financial institutions.	4.51	5.99	0	40
Number of Loans ≥ 2	Binary indicator = 1 if number of loans from financial institutions ≥ 2.	0.73	0.44	0	1
Age and Size of Firm					
Age of Firm (years)	Years firm has been in operation.	18.11	13.37	0.08	103
Assets ($millions)	Firm's total physical and financial assets.	4.87	13.5	0.001	191
Number of Employees	Total number of owner and non-owner employees.	57.6	76.2	1	486
Financial State					
Profits/Assets	Profits per asset.	0.69	2.22	−1.46	17.24
Cash/Assets	Cash per asset.	0.18	0.25	0	1
Leverage	Total Debt/Assets	0.61	1.25	0	11
Credit Quality					
Dun & Bradstreet Credit Score Indicator (in {1,2,3,4,5,6})	Indicates relative position of firm's D&B credit score in overall distribution of credit scores. 1=most risky; 6=least risky.	4.05	1.43	1	6
Bankruptcy in the past 7 years.	Binary indicator = 1 if firm has declared bankruptcy within past 7 years.	0.01	0.07	0	1
Judgments in the Past 3 Years.	Binary indicator = 1 if firm has had judgments rendered against it in past 3 years.	0.03	0.18	0	1
Delinquent Obligations in the Past 3 Years.	Binary indicator = 1 if firm has had any delinquent obligations in past 3 years.	0.18	0.38	0	1
Financial Statements Audited	Binary indicator = 1 if firm's financial statements were audited.	0.09	0.28	0	1
Loan Applications Always Approved Over the Past 3 Years.	Binary indicator = 1 if every loan application made by firm in past 3 years was approved.	0.97	0.18	0	1
Fear Denial of Loan Application	Binary indicator = 1 if ever in past 3 years firm did not apply for a loan out of fear of being denied (rejected).	0.11	0.31	0	1

Table 1 (continued). Definitions and Summary Statistics. Definitions and statistics for variables used in estimation. Data consist of 1,261 observations on 1,261 loans. Some observations are incomplete. Except for macroeconomic interest rates, all data come from the 2003 Survey of Small Business Finances.

Variable Name	Variable Definition	Mean	St.Dev	Min	Max
Organization and Ownership					
Firm is a Corporation	Binary indicator = 1 if the firm files taxes as a corporation.	0.77	0.42	0	1
Owners' Experience (years)	Average business experience of owners, weighted by ownership share.	23.00	10.91	0	65
Sales Radius Indicator (in {1,2,3,4,5,6,7})	Indicates radius of firm's market – where firm sells its goods or services. 1 = local city; 2 = county or MSA; 3 = state; 4 = region; 5 = throughout U.S.; 6=outside U.S.; 7=global (national and international).				
Geography					
Urban (in MSA)	Binary indicator = 1 if firm's headquarters are located in an MSA.	0.78	0.42	0	1
Census Division Indicator (in {1,2,3,...,9})	Indicates the census division where firm's headquarters is located. The U.S. has 9 official census divisions.				
Economy					
Banking Market HHI Indicator (in {1,2,3})	Indicator = 1 if banking deposit market HHI in (0,0.1); Indicator = 2 if HHI in (0.1,0.18); Indicator = 3 if HHI ≥ 0.18.	2.02	0.73	1	3
Prime Rate (percent)	Bank prime loan rate. Rate posted by a majority of top-25 (by domestic assets), insured, U.S.-chartered commercial banks.	4.24	0.31	4.00	5.25
Default Spread (percent)	Moody's seasoned Baa corp. bond yield minus 30-yr treas. constant maturity rate.	1.48	0.33	1.23	2.62
Term Spread (percent)	Constant maturity treasury rate for debt of same maturity as firm's loan minus 3-month treasury constant maturity rate.	1.20	1.05	0.00	4.44
Most-Recent Loan Characteristics					
Distance from Firm to Lender (miles)	Distance from firm's headquarters to lender's office where firm applied for loan.	45.74	193.7	0	2,500
Indicator of Type of Loan (in {1,2,...,7})	Type of loan. 1=new line of credit; 2=capital lease; 3=mortgage; 4=vehicle loan; 5=equipment loan; 6=other loan; 7=renewed line of credit.				
Indicator of Type of Financial Institution Used	Type of institution. 1=commercial bank; 2=savings bank; 3=S&L; 4=credit union; 5=finance co.; 6=insurance co.; 7=brokerage; 8=leasing co.; 9=mortgage co.				
Amount of Credit Granted ($millions)	Size of loan.	1.30	4.67	0.0003	83.80
Loan Maturity (years)	Length of time over which loan is to be repaid, as originally contracted.	3.58	4.73	0	36
Interest Rate is Fixed	Binary indicator = 1 if the interest rate on the loan is fixed.	0.42	0.49	0	1
Interest Rate (%)	Loan interest rate.	5.63	2.56	0	30
Indicator of Type of Collateral (if any) (in {1,2,3,4,5,6,7})	Type of collateral. 0 = no collateral; 1 = inventory or accounts receivable; 2 = equipment or vehicles; 3 = securities or deposits; 4 = business real estate; 5=personal real estate; 6=other personal assets; 7=other.				

Table 1 (continued). Definitions and Summary Statistics. Definitions and statistics for variables used in estimation. Data consist of 1,261 observations on 1,261 loans. Some observations are incomplete. Except for macroeconomic interest rates, all data come from the 2003 Survey of Small Business Finances.

Variable Name	Variable Definition
Reasons for Applying to Lender Reasons for Applying to Lender (42 indicators)	Binary indicators for each one of 42 possible lender attributes a firm can list as a reason for applying to the lender that supplied its most-recent loan. Each firm lists up to 3 reasons and can have up to 3 non-zero indicators. Reasons include: credit availability; previous loan; only institution to grant credit; turned down by other institutions; no or less collateral; SBA loan availability; lending policies or terms; credit availability contingent on use of other services; large loan capability; service availability (including credit card processing availability); quality of service; location or proximity; convenience or ease of use; reputation; small size of institution; large size of institution; size of institution (small/large not ascertainable); one-stop shopping (able to obtain multiple services at same institution); internet or electronic services; knowledge of industry; local bank; specializes in small business services; availability; good prices or terms; low fees or prices; high interest rate on savings or checking accounts; low interest rate or low loan origination fees; interest rate (high or low not ascertainable); captive finance (e.g. used financial institution owned by seller of purchased asset); friendly, knowledgeable employees or management; long-term relationship/prior relationship; reciprocity (institution does business with firm); primary bank; only institution in town or limited choice; institution or salesman solicited firm; original institution taken over by current one; diversification (convenient to use multiple institutions); minority ownership in institution; dissatisfaction with previous institution; other ascertainable reason; credit needed and no other response given; non-ascertainable response.
Industry Fixed Effects 2-Digit SIC Code Indicators (54 indicators)	Binary indicators for each one of 54 possible 2-digit SIC Industry Codes. Industries (codes in parentheses) include: oil and gas extraction (13); building construction general contractors and operative builders (15); heavy construction other than building construction contractors (16); construction special trade contractors (17); food and kindred products manufacturing (20); textile mill products (22); apparel and other finished products made from fabrics and similar materials (23); lumber and wood products, except furniture (24); furniture and fixtures (25); paper and allied products (26); printing, publishing, and allied industries (27); chemicals and allied products (28); rubber and miscellaneous plastics products (30); stone, clay, glass, and concrete products (32); primary metal industries (33); fabricated metal products, except machinery and transportation equipment (34); industrial and commercial machinery and computer equipment (35); electronic and other electrical equipment and components, except computer equipment (36); transportation equipment (37); measuring, analyzing, and controlling instruments; photographic, medical and optical goods; watches and clocks (38); miscellaneous manufacturing industries (39); local and suburban transit and interurban highway passenger transportation (41); motor freight transportation and warehousing (42); water transportation (44); transportation by air (45); transportation services (47); communications (48); electric, gas, and sanitary services (49); wholesale trade – durable goods (50); wholesale trade – non-durable goods (51); building materials, hardware, garden supply, and mobile home dealers (52); general merchandise stores (53); food stores (54); automotive dealers and gasoline service stations (55); apparel and accessory stores (56); home furniture, furnishings, and equipment stores (57); eating and drinking places (58); miscellaneous retail (59); insurance agents, brokers, and service (64); real estate (65); hotels, rooming houses, camps, and other lodging places (70); personal services (72); business services (73); automotive repair, services, and parking (75); miscellaneous repair services (76); motion pictures (78); amusement and recreation services (79); health services (80); legal services (81); educational services (82); social services (83); museums, galleries, and botanical and zoological gardens (84); engineering, accounting, research, management, and related services (87); miscellaneous services (89).

Table 2. Summary Statistics By Relationship Status. Statistics for select variables used in estimation by relationship status. Data consist of 1,261 observations on 1,261 loans. Some observations are incomplete. Except for macroeconomic interest rates, all data come from the 2003 Survey of Small Business Finances.

Variable	No Relationship (N=84)	Relationship Exists (N=1,177)	Difference
Dependent Variable			
Origination Cost of Most Recent Loan ($thousands)	8.43	2.78	5.65***
Treatment Variable			
No Relationship (New Lender)	1	0	1
Financial Relationships			
Max Relationship Length (years)	9.86	15.86	−6.00***
Number of Loans	4.12	4.53	−0.38
Number of Loans ≥ 2	0.74	0.73	0.01
Age and Size of Firm			
Age of Firm (years)	12.89	18.49	−5.59***
Assets ($millions)	5.54	4.82	0.72
Number of Employees	56.0	57.7	−1.71
Financial State			
Profits/Assets	0.62	0.70	−0.07
Cash/Assets	0.21	0.18	0.03
Leverage	0.74	0.60	0.14
Credit Quality			
Dun & Bradstreet Credit Score Indicator (in {1,2,3,4,5,6})	3.76	4.07	−0.31*
Bankruptcy in the past 7 years.	0	0.01	−0.01***
Judgments in the Past 3 Years.	0.01	0.03	−0.02
Delinquent Obligations in the Past 3 Years.	0.21	0.18	0.04
Financial Statements Audited	0.02	0.09	−0.07***
Loan Applications Always Approved Over the Past 3 Years.	0.92	0.97	−0.05*
Fear Denial of Loan Application	0.14	0.11	0.04
Organization and Ownership			
Firm is a Corporation	0.75	0.77	−0.02
Owners' Experience (years)	18.73	23.29	−4.56***
Geography			
Urban (in MSA)	0.86	0.77	0.09**
Economy			
Banking Market HHI Indicator (in {1,2,3})	1.92	2.02	−0.09
Prime Rate (percent)	4.20	4.24	−0.05
Default Spread (percent)	1.51	1.47	0.04
Term Spread (percent)	1.75	1.17	0.59***
Most-Recent Loan Characteristics			
Distance from Firm to Lender (miles)	163.75	36.31	126.44***
Amount of Credit Granted ($millions)	2.56	1.21	1.34
Loan Maturity (years)	6.24	3.41	2.83***
Interest Rate is Fixed	0.55	0.41	0.14**
Interest Rate (%)	5.91	5.61	0.30

Statistical significance of differences is calculated using t-tests with unequal variance.
Significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.

Table 3. Regression Using All Firms With Select Covariates. Table shows OLS regressions of small business loan origination costs on firms' selection to borrow from an institution with which they have no prior financial relationship. The dependent variable equals the total loan origination costs on charged on a firm's most recent loan. The key independent variable is a binary indicator of a firm's selection to borrow from a previously unused financial institution.

***Dependent Variable*: Origination Costs of Most-Recent Loan** ($thousands)

	(1)	(2)	(3)	(4)
No Relationship (New Lender)	5.65***	6.20***	6.34***	5.85***
	(2.90)	(3.41)	(3.45)	(3.17)
log Max Relationship Length		−0.58	−0.48	−0.25
		(1.39)	(1.14)	(0.63)
Number of Loans		−0.021***	−0.023***	−0.019***
		(2.62)	(2.75)	(2.88)
log Age of firm		−0.94**	−0.93**	−0.78*
		(2.15)	(2.07)	(1.71)
log Assets		1.71**	1.77***	0.44***
		(8.23)	(8.19)	(2.66)
Profits/Assets			0.058	0.078***
			(1.11)	(2.66)
Cash/Assets			0.029***	0.015***
			(5.37)	(2.92)
D&B Credit Score 26-50			−0.33	−0.69
			(0.28)	(0.75)
D&B Credit Score 51-75			−0.31	−0.74
			(0.31)	(0.98)
D&B Credit Score 76-90			−0.52	−1.07
			(0.49)	(1.28)
D&B Credit Score 91-100			−1.14	−1.95**
			(1.07)	(2.45)
log Amount of Credit Granted (Size of Loan)				2.19***
				(7.77)
log Maturity				1.10***
				(3.77)
Interest Rate (%)				0.35***
				(3.72)
Mortgage for Business Purposes				2.68**
				(2.03)
Vehicle Loan for Business Purposes				1.65**
				(2.39)
Constant	2.78***	−4.79***	−5.12***	−29.2***
	(10.63)	(4.07)	(3.16)	(7.87)
Number of Observations	1,261	1,251	1,248	1,179
R^2	0.02	0.14	0.14	0.25

Absolute values of heteroskedasticity-robust t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.

Table 4. Regression Using All Firms With All Covariates. Table shows OLS regressions of small business loan origination costs on firms' selection to borrow from an institution with which they have no prior financial relationship. The dependent variable equals the total loan origination costs on charged on a firm's most recent loan. The key independent variable is a binary indicator of a firm's selection to borrow from a previously unused financial institution.

***Dependent Variable*: Origination Costs of Most-Recent Loan** ($thousands)

	(1)	(2)	(3)	(4)
No Relationship (New Lender)	5.65***	6.18***	6.74***	5.83***
	(2.90)	(3.55)	(3.65)	(2.88)
Financial Relationships				
log Max Relationship Length		YES	YES	YES
Number of Loans; Number of Loans ≥ 2				
Num Applications to Get MR Loan ≥ 2				
Age and Size of Firm				
log Age of Firm; log Assets		YES	YES	YES
log Number Employees				
Financial State				
Profit/Assets;		YES	YES	YES
Cash/Assets;				
Leverage (Debt/Assets).				
Credit Quality				
D&B Credit Score Indicators;				
Bankruptcy in past 7 yrs;				
Judgments in past 3 yrs			YES	YES
Delinquent Obligations in past 3 years;				
Financial Statements Audited				
Loan Applications Always Approved				
Fear Denial of Loan Application				
Organization and Ownership				
Firm is a Corporation			YES	YES
log Owners' Experience				
Sales Radius Indicators				
Geography				
log Mean Distance to Financial Institutions			YES	YES
Urban (in MSA)				
Census Division Indicators				
Economy				
Banking Market HHI indicators				YES
Prime Rate; Default Spread; Term Spread;				
Most-Recent Loan Characteristics				
log Distance from Firm to Lender				
Indicators for Type of Loan				
Indicators for Type of Financial Institution Used				YES
log Amount of Credit Granted; log Maturity				
Interest Rate is Fixed; Interest Rate (%)				
Indicators for Type of Collateral (if any)				
Reasons for Applying to Lender				
Firm's Reasons for Applying for a Loan from the Financial				YES
Institution that Supplied the Firm's Most-Recent Loan				
(1 indicator for each possible reason)				
Industry Fixed Effects				YES
2-Digit SIC Code Indicators				
Number of Observations	1,261	1,251	1,248	1,178
R^2	0.02	0.15	0.23	0.41

Absolute values of heteroskedasticity-robust t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.

Table 5. Firms Among Which There is Less Potential for Endogenous Selection. Table shows OLS regressions of small business loan origination costs on firms' selection to borrow from an institution with which they have no prior financial relationship. The dependent variable equals the total loan origination costs on charged on a firm's most recent loan. The key independent variable is a binary indicator of a firm's selection to borrow from a previously unused financial institution.

***Dependent Variable*: Origination Costs of Most-Recent Loan** ($thousands)

	(1)	(2)	(3)	(4)
No Relationship (New Lender)	7.15**	7.98***	8.25**	6.14**
	(2.23)	(2.65)	(2.76)	(1.99)
Financial Relationships		YES	YES	YES
Age and Size of firm		YES	YES	YES
Financial State		YES	YES	YES
Credit Quality			YES	YES
Organization and Ownership			YES	YES
Geography			YES	YES
Economy				YES
Most-Recent Loan Characteristics				YES
Reasons for Applying to Lender				YES
Industry Fixed Effects				YES
Number of Observations	588	586	584	558
R^2	0.03	0.16	0.26	0.48

Absolute values of heteroskedasticity-robust t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *$p<0.10$; **$p<0.05$; ***$p<0.01$.

Table 6. Natural Quasi-Experiment Regression. Table shows analysis of firms that select lenders based on proximity. Table shows OLS regressions of small business loan origination costs on firms' selection to borrow from an institution with which they have no prior financial relationship. The dependent variable equals the total loan origination costs on charged on a firm's most recent loan. The key independent variable is a binary indicator of a firm's selection to borrow from a previously unused financial institution.

***Dependent Variable*: Origination Costs of Most-Recent Loan** ($thousands)

	reason is distance			reason is distance and firm is less constrained in selection of lender		
	(1)	(2)	(3)	(4)	(5)	(6)
No Relationship (New Lender)	13.23**			13.69**		
	(2.07)			(2.09)		
Reason is NOT Distance		omitted	omitted			
Reason is NOT Distance × No Relationship		5.36***	5.31**			
		(2.68)	(2.60)			
Reason is Distance		−1.54***	−1.66			
		(3.16)	(1.52)			
Reason is Distance × No Relationship		13.23**	10.64**			
		(2.11)	(2.12)			
NOT (Dist and Less Constr)					omitted	omitted
NOT (Dist and Less Constr) × No Relationship					5.39***	5.28**
					(2.69)	(2.55)
Distance and Less Constrained					−1.98***	1.16
					(4.75)	(0.57)
Distance and Less Constr × No Relationship					13.69**	10.14**
					(2.18)	(2.03)
Financial Relationships			YES			YES
Age and Size of Firm			YES			YES
Financial State			YES			YES
Credit Quality			YES			YES
Organization and Ownership			YES			YES
Geography			YES			YES
Economy			YES			YES
Most-Recent Loan Characteristics			YES			YES
Reasons for Applying to Lender			YES			YES
Industry Fixed Effects			YES			YES
Number of Observations	47	1,261	1,178	23	1,261	1,178
R^2	0.42	0.02	0.42	0.55	0.02	0.42

Absolute values of heteroskedasticity-robust t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.

Table 7. AET. Table shows AET estimation of the relationship between loan origination costs and firms' selection to borrow from an institution with which they have no prior financial relationship. Estimation follows the bivariate maximum-likelihood methodology of Altonji, Elder, and Taber (2002, 2005). The dependent variable equals the total loan origination costs on charged on a firm's most recent loan. The treatment variable is a binary indicator of a firm's selection to borrow from a previously unused financial institution.

***Outcome Variable*: Origination Costs of Most-Recent Loan** ($thousands)
***Treatment Variable*: No Relationship (New Lender)**

	ρ fixed at $\rho = 0$		ρ estimated	
	select controls included	all controls included	select controls included	all controls included
	(1)	(2)	(3)	(4)
$\hat{\rho} = \text{corr}(\varepsilon, u)$			−0.039 {0.18} [0.21]	−0.074 {0.11} [0.66]
No Relationship (New Lender)	6.42*** (3.50)	5.74*** (3.17)	6.98** (3.87) [2.22]	6.04*** (3.43) [3.10]
Financial Relationships	YES	YES	YES	YES
Age and Size of Firm	YES	YES	YES	YES
Financial State	YES	YES	YES	YES
Credit Quality	YES	YES	YES	YES
Organization and Ownership	YES	YES	YES	YES
Geography	YES	YES	YES	YES
Economy		YES		YES
Most-Recent Loan Characteristics		YES		YES
Reasons for Applying to Lender		YES		YES
Industry Fixed Effects		YES		YES
Number of Observations	1,248	1,178	1,248	1,178
R^2	0.23	0.41	0.23	0.41

Standard errors calculated using nonparametric bootstrap shown in curly brackets, { }.
Absolute values of t-statistics obtained from bootstrapped standard errors shown in square brackets, [ ].
Heteroskedasticity-robust asymptotic ML t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.

Table 8. Loan Interest Rates, Price Sensitivity, and Relationships. Table shows analysis of firms that select lenders based on sensitivity to loan prices. Table shows OLS regressions of loan interest rates on firms' price sensitivity state and firms' selection to borrow from an institution with which they have no prior financial relationship. The dependent variable equals the loan interest rates charged on a firm's most recent loan. The key independent variables are a binary indicator of a firm's selection to borrow from a previously unused financial institution and a binary indicator of a firm's self-reported sensitivity to loan prices.

***Dependent Variable*: Interest Rate on Most-Recent Loan** (percent)					
	(1)	(2)	(3)	(4)	(5)
Reason is Price	−0.74***		−0.63***	−0.61***	−0.62***
	(4.29)		(3.83)	(3.52)	(3.18)
No Relationship (New Lender)		0.30	−0.02	−0.11	−0.42
		(0.76)	(0.06)	(0.28)	(0.98)
Financial Relationships			YES	YES	YES
Age and Size of firm			YES	YES	YES
Financial State			YES	YES	YES
Credit Quality				YES	YES
Organization and Ownership				YES	YES
Geography				YES	YES
Economy					YES
Most-Recent Loan Characteristics					YES
Industry Fixed Effects					YES
Number of Observations	1,261	1,261	1,251	1,248	1,178
R^2	0.013	0.001	0.18	0.24	0.35

Absolute values of heteroskedasticity-robust t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.

Table 9. Loan Origination Costs, Price Sensitivity, and Relationships. Table shows analysis of firms that select lenders based on sensitivity to loan prices. Table shows OLS regressions of loan interest rates on firms' price sensitivity state and firms' selection to borrow from an institution with which they have no prior financial relationship. The dependent variable equals the total loan origination costs charged on a firm's most recent loan. The key independent variables are a binary indicator of a firm's selection to borrow from a previously unused financial institution and a binary indicator of a firm's self-reported sensitivity to loan prices.

***Dependent Variable*: Origination Costs of Most-Recent Loan** ($thousands)					
	(1)	(2)	(3)	(4)	(5)
Reason is Price	1.12 (1.42)		0.46 (0.61)	0.24 (0.29)	−0.27 (0.32)
No Relationship (New Lender)		5.65*** (2.90)	6.09*** (3.35)	6.69*** (3.61)	5.63*** (2.89)
Financial Relationships			YES	YES	YES
Age and Size of firm			YES	YES	YES
Financial State			YES	YES	YES
Credit Quality				YES	YES
Organization and Ownership				YES	YES
Geography				YES	YES
Economy					YES
Most-Recent Loan Characteristics					YES
Industry Fixed Effects					YES
Number of Observations	1,261	1,261	1,251	1,248	1,178
R^2	0.002	0.02	0.15	0.22	0.39

Absolute values of heteroskedasticity-robust t-statistics shown in parentheses.
Statistical significance indicated by asterisks: *p<0.10; **p<0.05; ***p<0.01.