w o r k i n g
p a p e r

06 16

Bank Branch Presence and Access to Credit in Low-to-Moderate Income Neighborhoods

by O. Emre Ergungor

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 06-16 December 2006

Bank Branch Presence and Access to Credit in Low-to-Moderate Income Neighborhoods

By O. Emre Ergungor

Banks specialize in lending to informationally opaque borrowers by collecting soft information about them. Some researchers claim that this process requires a physical presence in the market to lower information collection costs. I provide evidence in support of this argument in the mortgage market for low-income borrowers. Mortgage originations increase and interest spreads decline when there is a bank branch located in a low-to-moderate income neighborhood.

Keywords: Access to credit, branch proximity

O. Emre Ergungor is an economist at the Federal Reserve Bank of Cleveland and can be reached at ozgur.e.ergungor@clev.frb.org or 216-579-3004.

1. Introduction

Access to credit is vital for the poor. Credit insures individuals against liquidity shocks, prevents unnecessary liquidation of illiquid investments, and channels savings from unproductive liquid assets toward investments in productive capital (Bencivenga and Smith, 1991). This process helps the poor because the ability to invest in productive assets and the associated increase in wealth encourage investment in human capital by the young generations, increasing the productivity of the poor and boosting their quality of life (Galor and Zeira, 1993; Wolfensohn and Bourguignon, 2004).

However, informational frictions could make credit prohibitively expensive or outright rationed (à la Stiglitz and Weiss, 1981) in low-income areas where credit histories are tainted by past problems or simply non-existent.[1] In this paper, I investigate whether banks play a role in alleviating these frictions by collecting soft information about their customers (i.e. information that is difficult to express in hard numbers, such as honesty and diligence) and making credit more accessible. One way to determine whether soft information is playing any role in the lending market is to examine how the distance between the lender and the borrower affects the availability and terms of credit (Petersen and Rajan, 1994; Degryse and Ongena, 2005; Hauswald and Marquez, 2006; Agarwal and Hauswald, 2006). So, I ask the following question in this paper: How does access to a bank branch affect the availability and cost of credit in low-income neighborhoods?

Using mortgage data disclosed by lenders under the Home Mortgage Disclosure Act (HMDA) of 1975 and reported by the Federal Financial Institutions Examination Council, I find a strong *positive* relationship between bank branch presence in low-income neighborhoods (a measure I develop to capture the number of branches inside and around a neighborhood) and mortgage originations and a strong *negative* relationship between branch presence and mortgage spreads over maturity-matched Treasury securities.

[1] Estimates for these types of consumers vary between 10 and 22 million households. "Innovations in Personal Finance for the Unbanked: Emerging Practices from the Field", Fannie Mae Foundation Case Studies, 2003.

I also find that the favorable effects of branch presence get stronger as the branch gets closer to the neighborhood. These findings reinforce earlier evidence presented by Petersen and Rajan (1994), which suggests that in the small-business-lending market, relationships are associated with greater availability of credit. By contrast, I find that branch presence is not correlated with mortgage availability in high-income neighborhoods, where borrowers are more likely to qualify for credit-scored mortgages.

This paper makes two important contributions to the literature. First, it complements studies that seek to identify the role played by financial intermediaries in alleviating poverty. Cross-country studies have shown that financial development alleviates poverty and reduces income inequality (Beck, Demirguc-Kunt, and Levine, 2004). Yet the mechanism through which financial development plays its favorable role is unclear. The mechanisms proposed by studies of microlending and proven effective ---peer monitoring, mutual insurance, etc. --- have not always been successfully replicated in other countries (MkNelly and Dunford, 1998, 1999; Conlin, 1999).[2] Even when they are successful, the reach of these programs remains tiny relative to the population (less than 2 percent of the population is served in developing countries), and their survival is mostly dependent on government or donor subsidies (Honohan, 2004a). My contribution is that I propose a mechanism---relationship lending---that does not suffer from any of these encumbrances. It is managed by mainstream financial intermediaries and it has been shown to alleviate information frictions in the small-business loan market, where borrowers' informational opacity is a well-known problem.

The second contribution of the paper is that it expands the relationship-lending literature. This literature has focused exclusively on the funding needs of small businesses and has attempted to show how relationships between banks and borrowers ease informational frictions in the lending market and thus allow greater access to credit. However, the evidence has not provided an unqualified support for the relationship

---

[2] Microloans are small loans (as small as $75) to the poor to help them finance self-employment activities. See Pulley, 1989; Adams and von Pischke, 1992; Christen et al., 1995; Goetz and Gupta, 1996; Jain, 1996; Coleman, 1999; Morduch, 1999; Wydick, 1999; Coleman, 2002; Matin, Hulme, and Rutherford, 2002; Anderson, Baland, and Moene, 2003; Beegle, Dehejia, and Gatti, 2003, to name just a few. Theoretical underpinnings of this literature can be found in Stiglitz, 1990; Varian, 1990; Besley, Coate, and Loury, 1993; Besley and Coate, 1995; Ghosh and Ray, 1996; Ghatak, 1999; Armendáriz de Aghion and Morduch, 2004.

lending theory (more on this in the next section). Mortgages, on the other hand, have traditionally not been viewed as relationship loans (Stein, 2002, p. 1892). Despite the heavy use of credit-score based models in mortgage originations, there still seems to be a special group of borrowers in low-income neighborhoods who need the assistance of a lender willing to invest in the collection of soft information; they constitute an interesting laboratory where the implications of relationship lending can be tested.

The remainder of the paper is organized as follows. The next section explains how relationships ease informational frictions and presents the hypotheses tested in later sections. Section 3 describes the data. Section 4 explains the method. Section 5 presents the results. Section 6 concludes.

2. Background and Hypotheses

The claim that branch presence in low-income neighborhoods makes credit more accessible to the poor builds on the relationship-lending theory (Sharpe, 1990; Rajan, 1992; von Thadden, 2002). As Stiglitz and Weiss (1981) have shown, the lack of information about a borrower's credit quality leads to credit rationing due to adverse selection. The market response to potential credit rationing has been credit-scoring and relationship lending. Credit scoring allows lenders to judge the riskiness of a borrower based on incomplete information about the borrower, so long as a particular subset of borrower characteristics can be observed at a relatively low cost. Relationship lending, on the other hand, relies on soft information about borrowers that is observed through a lender's interactions with the borrower through time. The strength of the relationship will be a function of time and the diversity of these interactions (Berger and Udell, 1995; Degryse and Van Cayseele, 2000). A bank's advantage in relationship lending is that it can interact with a borrower on both sides of the balance sheet -- through lending and deposit products -- and in such an environment, it can learn about the borrower's quality over time. While relationship development is a costly process, the bank can recoup its initial investment by exploiting its private information as the borrower graduates into more profitable product lines (Berger and Udell, 1996). Competitors know that attempts to lure away customers from the relationship lender subject them to a lemons problem, as the relationship bank will undercut them in the pricing of credit only in those cases where it is profitable to do so (Sharpe 1990).

The types of information that relationship lending depends on can only be reliably collected and processed at a local level. In other words, relationship lending requires a physical presence in the market. In this setting, the distance between the lender and borrower is important because proximity lowers the cost of collecting soft information (e.g., extra communication costs). As a result, loan applicants close to their lenders are more likely to be approved and less likely to default (Petersen and Rajan, 2002; Brevoort and Hannan, 2004; DeYoung et al 2006).

The relationship theory has never been applied to consumer lending. The focus has been instead on small business loans. The choice of small businesses as a focal point is, to some extent, a consequence of the time period when this literature flourished. One major policy concern in the early 1990s was whether small businesses would starve for credit as small banks disappeared as a result of the merger wave of 1990s. In response, the U.S. Justice Department reviewed the antitrust implications of bank-merger applications under the assertion that local small business loan markets must be preserved and a physical presence in a market is necessary to adequately serve that market. It viewed the lack of physical facilities as prima facie evidence a bank is not serving a community. The concern was rooted in the long-held belief that small banks are crucial in lending to small businesses because these banks specialize in character loans rather than (or in addition to) arm's-length loans.[3]

Given the policy concerns of the time, it was perhaps only natural for economists to test the implications of relationship-lending theory within the context of small banks lending to small businesses (Berger, Saunders, Scalise, and Udell, 1998; Peek and Rosengren, 1998, Strahan and Weston, 1998 to name a few). The facts that emerged from these studies are less than an unqualified support for the theory. Although the evidence supports the view that small banks specialize in lending to informationally "difficult" borrowers (Berger et al., 2005; Cole et al., 2004), small firms' access to credit is not

[3] Large banks and other non-bank intermediaries, however, specialize in *cookie-cutter* loans, which are approved or denied based on a computer model (a credit score) and leave little wiggle room to the

impaired in the long run when there are fewer small banks in the area. Small firms in areas with few small banks are no more credit-constrained than firms in areas with many small banks (Jayaratne and Wolken, 1999).

The reason for the ambiguous findings may be the way small business lending is measured. First, a business owner with good personal credit can take a consumer loan and use it to finance his business. This loan is not classified as a small business loan despite its final use. When a small bank disappears, a consumer loan may substitute for a relationship loan and make it difficult to estimate the value of the bank-small business relationship. Second, small businesses are not a homogenous group. The U.S. Small Business Administration classifies an enterprise as a "small business" if its assets are less than a pre-specified amount or if it has fewer than a pre-specified number of employees---depending on the industry. Under these guidelines, a mom-and-pop grocery store, a $28 million firm in highway construction or a 1,500-employee firm in aircraft research can all be considered small businesses. Finally, if one thinks of a small business loan as a loan to a company that does not have access to the public bond market, then even loans as large as $100 million may fit the description of a "relationship" loan. So, the ambiguity in the definition of a small business may be responsible for the conflicting/ambiguous results in the literature.

One solution to these problems may be to step back and test the theory from a different perspective. Mortgage loans in low-income areas are a good place to start.[4] The benefit of examining mortgage loans is that data aggregation problems are less of a concern. The definition of a "low-income consumer" is less ambiguous than the definition of a small business. Also, consumers do not substitute mortgages for other types of loan, a possible strategy with small business loans.

By applying the tenets of relationship lending to consumer lending, I conjecture that lenders' interactions with households in low and moderate income neighborhoods

---

judgment of a loan officer. Borrowers choose among a pre-determined set of contracts, and loan terms do not adjust to borrowers' specific needs. (Cole, Goldberg, and White, 2004)

[4] One would imagine that information-driven agency problems in lending markets would be more severe in local credit markets consisting of low and moderate income households.

provide them with an informational advantage over other lenders.[5] These interactions take place through the delivery of services that are targeted to the poor, such as financial education and counseling services, "second-chance" checking and savings accounts, and low-cost alternatives to payday loans. A physical presence in the market is required to reach customers in these neighborhoods and glean information, such as how many checks a customer bounces every month, how many weeks the customer can go without using a payday loan or how often the customer has to renegotiate the repayment schedule. As the bank gets to know its customers more intimately, those who show promise may graduate into higher-margin, more sophisticated services, such as mortgages and auto loans. In other words, the bank gains from this relationship not necessarily through the repeat sales of the same product (as might be the case with small-business loans) but through the cross-selling of multiple products.

2.1 Testable Hypotheses

Consider a mortgage market comprised of borrowers with unobservable risk characteristics. Borrowers' expected return from their housing investment is such that, as in Stiglitz and Weiss (1981), loan demand declines at high interest rates as low-risk customers drop out of the market. With the mix of the applicant pool worsening, the supply also declines at high interest rates because the expected return to the bank per dollar loaned goes down. If the lender's profit is maximized at an interest rate below the market-clearing rate, demand is greater than supply in equilibrium.

The main hypothesis of the paper is based on the premise that reducing information barriers through relationship lending would mitigate the worsening of the applicant pool at high interest rates and enhance credit availability.

Hypothesis 1: *Access to credit, measured by the dollar amount of mortgages originated per household, will increase with greater access to bank branches, as measured by proximity to bank branches, in low-income neighborhoods.*

A related question is what happens to the price of credit in a market with increasing access to bank branches. Earlier studies in the small-business loan market

[5] Mester, Nakamura, and Renault (1998) show that tracking a firm's checking account balances may provide the bank useful information for detecting problems early on. Avery et al. (1999) find that banks

have found evidence supporting the predictions of the spatial-price-discrimination (SPD) models; i.e., the borrowing costs increase as the lender gets closer to the borrower but decline if there are other competing lenders nearby (Degryse and Ongena, 2005; Agarwal and Hauswald, 2006). The intuition is that the lender's proprietary information allows it to earn monopoly rents from its dealings with the borrower. As the distance increases, the informational advantage wanes. In other words, as the borrower gets farther and farther from the informed lender, the loan rates approach the competitive rate as the ability of the lender to collect the information and outcompete other lenders in the market disappears. Furthermore, as the uninformed competing lenders get closer to the borrower, the competitive rate itself declines due to declining transportation costs (Lederer and Hurter, 1986; Hauswald and Marquez, 2006).

How well this theory applies to the mortgage market for low-income borrowers is not obvious. There are at least two factors that set the low-income mortgage market apart. The first factor applies to the mortgage market in general, the second is low-income market specific.

The first factor is that arm's-length lending in the mortgage market does not depend on transportation costs. Transportation costs capture the borrower's difficulty in identifying the location and the loan-terms of the lenders in the area. Arm's-length mortgage market, however, is national. Those mortgages can be originated by distant lenders who base their decisions on the downpayment amount and the credit score (Chomsisengphet and Pennington-Cross, 2006). As a result, any potential borrower, irrespective of his income, can apply for a mortgage online or by responding to a mail offer. Because the arm's length mortgage market is national, the distance of uninformed lenders to the borrowers should be irrelevant to the mortgage rate.

The second factor is that Stiglitz-Weiss style credit rationing is not a problem considered in the SPD-based small-business lending literature; all borrowers of all types get credit at some price. In contrast, credit rationing is an important characteristic of the low-income mortgage market; uninformed lenders will never satisfy all the demand at any price.

---

learn about the neighborhood they are in as they process more mortgage applications.

How do these factors affect the predictions on the SPD models? First, unlike the arm's-length market, relationship lending is local. Therefore, I suppose that closer distance and increased number of bank branches have a detectable impact on the mortgage rates not necessarily because of higher competition by uninformed lenders but because closeness (lower transportation costs) provides a greater opportunity to form relationships. Second, because of credit rationing by the uninformed lenders, the competitive threat to the monopoly rents of the informed lender could be much smaller than anticipated by the SPD models. To the extent that the informed lender's monopoly rents are not challenged *ex post*, we may observe an increase in interest rates (and a decline in quantity) with increasing distance to reflect the lender's increasing risk due to deteriorating information quality.

The second hypothesis of the paper captures this idea.

Hypothesis 2: *Price of credit will decline with greater access to bank branches in low-income neighborhoods.*

3. Data Description

I analyze the effect of branch presence on access to mortgages in Ohio's low-income neighborhoods.[6] My definition of a neighborhood is a census tract. Census tracts are designed by the Census Bureau to be relatively homogeneous units in terms of population (about 4,000 inhabitants), population characteristics, economic status, and living conditions. All information related to census tract characteristics comes from the 1990 and 2000 Decennial Census. Information related to mortgages originated in a census tract comes from the 2004 HMDA Loan Application Register (LAR) data. Branch addresses that I use to determine the branch location come from the FDIC's Summary of Deposits file. In the remainder of this section, I will describe these data in greater detail.

3.1 Census Data

As described in the next section, I use the changes in census tract characteristics from 1990 to 2000 in explaining the changes in the local banking market. One complication is that census tract boundaries changed from the 1990 Census to the 2000

Census. In 1990, there were 2,862 census tracts in Ohio; in 2000, there were 2,941 tracts. The change in boundaries comes in many forms. Tracts split, tracts merge, and tracts split and merge with split-off sections of other tracts. To control for changes in tract characteristics, I need tract boundaries that are stable from 1990 to 2000.

To achieve stable boundaries, I use the strategy summarized in Figure 1. Using the Census Bureau's Census Tract Relationship Files, I merge the tracts that went through a boundary change until I cover the smallest area that has not changed its shape from 1990 to 2000. For example, in Figure 1, Census-1990 tracts 1, 2, 3, and 4 merge to form the new tract 1, while the Census-2000 tracts 1 and 2 merge to create the same new tract 1 in 2000. The characteristics of the new tract are obtained by combining the characteristics of the original tracts either by a simple addition (population, housing units, etc.) or a weighted average (median incomes weighted by population, median home values weighted by housing units, etc.). The redrawing of boundaries leaves me with 2,062 new census tract-based markets, 1,665 of them are the same as the Census Bureau's tracts; i.e., their shape remains constant from 1990 to 2000. The remainder, 387 areas, are an amalgamation of multiple tracts, of which 60 are an amalgamation of 4 tracts or more. A total of 39 census tracts are eliminated from the dataset because there is no resident population or no owner-occupied property (business districts, airports, etc.).

After this clean-up, I divide my sample of 2,023 local markets into two groups by census tract median income. HMDA classifies census tracts that are below the national median income as low-moderate income neighborhoods. Since I work with Ohio neighborhoods, I use the median income in Ohio as the cutoff point. By this standard, census tracts with median incomes below $40,956 are classified as low-moderate income census tracts (1,289 tracts). Those with median incomes above $40,956 are in the middle-high income category (734 tracts).

---

[6] I chose Ohio because the Federal Reserve Bank of Cleveland Community Affairs Department granted me access to their geocoding software, which I used to locate bank branches; more on this later.

3.2 Branch Presence

The FDIC's Summary of Deposits file provides the branch addresses of every FDIC-insured institution in the country. There were 3,886 bank branches in Ohio in 1999 and 3,963 in 2004. Using CRAWiz, a geocoding software package, each address is matched to a latitude and longitude. About 92 percent of the addresses match automatically; because of spelling errors or incomplete addresses, the rest must be matched manually. For those, I search for the correct address on the Internet and replace the old address with the new one in CRAWiz. If that fails, there are a few other alternatives. If the address is an intersection, I can point to the intersection on the CRAWiz map, and the software will use the latitude and longitude of that point. If there is ambiguity about the directional qualifier (e.g., North vs. South Main Street), I use Google satellite pictures to determine where the branch is located; for example, if 123 North Main Street is a residence and 123 South Main Street is a business building, the branch is in the business building. Using this method, I determine the location of every FDIC-insured institution branch in the state.

To obtain a measure of branch presence in a census tract, I determine the distance of each branch to the census tract centroid using the Haversine Formula (Sinnott, 1984).[7] Then, I take all the branches within 10 miles of the centroid and calculate the branch-access variable as:

$$BAccess_i = \ln\left(1 + \sum_{k=1}^{n_b} \frac{1}{D_{i,k}}\right) \tag{1}$$

where $BAccess_i$ is the branch access variable for census tract i, and n_b is the number of branches within 10-mile radius of the centroid of census tract i. $D_{i,k}$ is the distance of branch k to the centroid of census tract i. In accordance with the relationship literature, this construction assumes that the farther the branch is from the centroid, the less likely it is to improve the accessibility of banking services in the census tract.

[7] The centroid is conceptually similar to the center of gravity of a 3-D object, except that it applies to a two dimensional shape.

Four important issues about this variable are worth mentioning. First, this measure is better than counting only the branches inside a census tract because in urban areas, one could miss a branch across a street if the tract boundary is the street. Including all branches within a certain distance to the tract solves this problem. Second, the implicit assumption is that branches farther than 10 miles have no effect on branch access. I will address this issue in robustness checks by reducing and increasing this radius. Third, measuring the distance of the branch to the tract centroid will be misleading in rural areas where the census tracts are very large and populated areas are in a little corner of the tract (e.g., Southeast Ohio). So the inverse distance of the branch to the centroid will underestimate the level of branch access in areas where the centroid is in an uninhabited area. I will address this issue in robustness checks by discarding sparsely populated areas. Fourth, if the branch is located exactly over the census tract centroid, *BAccess* will go to infinity. In my sample, there are very few branches that are closer than 1 mile to the centroid and none of them is closer than 0.02 miles. But, one could conceivably get extremely large *BAccess* values. I investigate this issue further in robustness checks by winsorizing *BAccess*, and alternatively, by redefining it as

$$BAccess_i' = \ln\left(1 + \sum_{k=1}^{n_b} \frac{1}{\max\left(1, D_{i,k}\right)}\right) \tag{1'}$$

which treats branches that are closer than 1-mile as if they were at exactly 1-mile.

3.3 HMDA

Under the Home Mortgage Disclosure Act of 1975, depository and non-depository financial institutions report all mortgage applications they receive in each census tract by disclosing the loan applicant's income, race, gender, and the loan's amount, whether it is FHA or VA-insured, whether it was kept on the originator's balance sheet or sold, the originator's identification number, whether the application was approved or denied and if it is denied, the reason for denial, and starting for the first time with the 2004 data (reported in 2005), the *spread* of the loan price over the rate of a Treasury security of comparable maturity at the time of origination *if* the spread exceeds 3%.[8] The loan price includes the interest rate as well as points, fees and premiums for private mortgage insurance.

I exclude refinancings and mortgages to purchase renter-occupied property from the analysis. I also exclude the data on mortgages purchased by financial institutions that were originated at an earlier time. This prevents double-counting, once at origination and once at the time of sale. After this clean-up, I am left with 246,327 mortgage originations (>$17.3 billion) in 2000, 91,517 of them (>$6.4 billion) in low-income neighborhoods and 266,516 home mortgages (>$22.9 billion) in 2004, 100,666 of them (>$8.7 billion) in low-income neighborhoods.

HMDA presents many data challenges. There are a few pieces of information one would like to have about these loans but is left wanting. The maturity and the loan-to-value ratio of the mortgage are unknown, the borrower's credit score and downpayment amount are unknown, as are other important terms of the loan such as whether the loan has a fixed or adjustable rate or whether it is a full or no-doc loan. Finally, because the loan price is reported only if its spread is 3 percentage points above the Treasury rate, some mortgages are reported with a zero interest rate; 69 census tracts have no price reported (45 of them low-moderate income) because *all* spreads are below 3%.[9]

Some of the problems of HMDA, such as the missing maturity and loan-to-value ratio, cannot be fixed. However, their impact will be limited in low-income communities. While affordable housing programs targeted to low-income neighborhoods, such as the Fannie Mae Community Home Buyer Program, Fannie 97 and FHA-insured mortgages, allow 15- or 30-year mortgages, it is safe to assume that cash-strapped individuals will be more inclined to take the 30-year mortgage (lower monthly payments make it easier to qualify for a 30-year mortgage than a shorter-term mortgage). Whether the mortgage has a fixed or adjustable rate may be unknown but this is again less of a problem in low-income areas relative to high-income areas because

[8] There are some exemptions to reporting requirements based on an institution's asset size or the size of its mortgage lending business. However, the reporting threshold is low enough that HMDA represents an accurate picture of the local lending market.
[9] It may seem surprising at first that there are more low-income census tracts with all spreads less than 3% than high-income census tracts. This is because almost two-thirds of Ohio's census tracts are low-income.

affordable housing programs require a fixed-rate mortgage. The missing downpayment data is also not too problematic because downpayment requirements in low-income areas are very low, varying between 0% and 5% (3% for FHA loans).

Finally, the price data can be improved upon. In order to estimate the local mortgage interest rate, I fit a lognormal distribution over the reported spread data in *each* census tract, assuming that the distribution is left-censored at 3%. I accomplish this by estimating a censored regression model including only an intercept on the right-hand side. The intercept is the mean of the uncensored distribution, which I use in the analysis.

4. Method

The quantity and price of loans in a local market must be simultaneously determined. I estimate the impact of bank branch presence in 1999, *BAccess99*, on mortgage originations---the dollar amount of mortgage originations per household in 2004, *Originate04*---and spreads, *Spread04*, by estimating the following system with GMM (to preclude heteroscedasticity problems) in low-income census tracts:

$$\begin{aligned} Originate04 &= f\,(BAccess99,\ Spread04,\ Retired00,\ \boldsymbol{X_1}) + \varepsilon_O \\ Spread04 &= f\,(BAccess99,\ Originate04,\ LenderCost00,\ \boldsymbol{X_1}) + \varepsilon_S \qquad (2) \end{aligned}$$

Identification is clearly an issue. I identify the price, *Spread04*, by the share of retired individuals in the census tract population, *Retired00*. Retired households in a low-income neighborhood are unlikely to be significant participants in the home-purchase market. In the Census data, *Retired00* also captures the population in retirement communities, who do not use mortgages. Because this population is at best a tiny portion of the market, the spreads on the loans originated should not depend on its share in the total population; however, originations should decline with its increased presence.

I identify the quantity of the loans originated, *Originate04*, by the cost of interest-bearing liabilities (interest expense to interest-bearing liabilities) of the highest-cost bank in the market, *LenderCost00*.[10] I assume that the mortgage rates in the market are

[10] Using the average interest cost of all lenders does not affect the results.

sufficiently high to make the highest-cost lender viable. Therefore, mortgage spreads should be positively correlated with *LenderCost00*. Originations---or mortgage demand---will depend on banks' cost structure indirectly, only to the extent that the costs affect the spread.

The system also contains a rich set of control variables, denoted by $\boldsymbol{X_1}$, which comprises of demographic, mortgage-market-specific, and banking-market-specific factors.

4.1 Demographic Factors

ChildInFamily00 is the share of children in the census tract who live in a two-parent household. Because single-parent households may be more cash-constrained, *ChildInFamily00* may be positively associated with originations and negatively associated with spreads.

HighSchool00 is the share of the population over 25 years of age whose educational achievement is a high school diploma or less. *HighSchool00* may be negatively associated with originations and positively associated with spreads because educational achievement can be an indicator of job opportunities.

Income90 and *Income00* are the natural log of the median income in the census tract in 1990 and 2000, respectively.

Jobs00 is the employment rate in the census tract.

Manufacturing00 is the share of manufacturing jobs in total jobs in the census tract in 2000. This variable captures the effect of the employer mix on the availability of mortgages under the assumption that manufacturing jobs could be correlated with the perceived stability of employment in the area.

Population90 and *Population00* are the natural log of the population of the census tract in 1990 and 2000.

Race00 is the share of the African-American population in the total population.

RuralPop00 is the share of the population that is classified as rural by the Census Bureau. Population turnover may be slower in rural areas, which may impact originations. Spreads may be higher in these areas because of the dependence of the local economies on farming, a volatile sector of the economy.

ShortCommute00 is the share of the employed population that works at home or commutes less than 30 minutes to work. *ShortCommute00* is a measure of access to alternative banking markets. Its effect on access to mortgages can go either way. On one hand, access to financial services in distant banking markets may increase competition in the local market. As monopoly rents disappear, mortgages may become available at greater quantity and lower cost. On the other hand, more competition in the home market may weaken relationships and have a negative impact on mortgage availability in low-income neighborhoods (Petersen and Rajan, 1995).

4.2 Mortgage-Market-Specific Factors

Applications00 is the number of mortgage applications per owner-occupied housing units. It is a proxy for population turnover in the area. If an exogenous factor that I am not explicitly controlling for is creating demand for mortgages, its effect will be captured in *Applications00*.

BankShare00 is the share of banks in mortgage originations in the census tract. This variable controls for the presence of institutions that are not insured by the FDIC (credit unions and non-bank mortgage lenders).

BorrowerIncome00 is the natural log of the average income of mortgage applicants who were approved for a mortgage. Note that mortgage terms are likely to depend both on neighborhood characteristics and borrower characteristics. While *BorrowerIncome00* captures the borrower characteristics, *Income90* and *Income00* capture the neighborhood characteristics.

CoApplicant00 is a share of loan applications with a co-applicant.

Conventional00 is the share of mortgages that have been originated with at least 25% equity and without any government insurance.

CreditProblem00 is the share of mortgage applicants denied credit because of poor credit histories in 2000. Even though I do not have any information on the credit quality of each borrower, I use *CreditProblem00* as a proxy for credit risk at local market level.

FHA00 is the share of mortgages that are insured by the FHA.

HomeValue90 and *HomeValue00* are the natural log of the median house price in the census tract in 1990 and 2000.

Institutions00 is the natural log of one plus the *number* of institutions that originated mortgages in the census tracts, including all depositories, and non-bank lenders.

Originate00 is the initial value of *Originate04* in 2000.

OOHousing00 is the share of owner-occupied housing in the total housing stock; the more owner-occupied housing there is, the greater the originations will be.

UnsoldLoan00 is the share of mortgages originated but not sold (kept on the balance sheet of the originator). If these mortgages are not sold because the verifiable information about the borrowers are unacceptable to buyers who base their decisions on heuristics, then *UnsoldLoan00* is a measure of the market's opacity.

4.3 Banking-Market-Specific Factors

Deposits99 is total deposits in bank branches within 10 miles of the census tract centroid, weighted by the inverse of their distance to the centroid per household, per dollar of income. *Deposits99* captures the savings rate and the level of financial development in the census tract. Keeping all other factors constant, higher values of *Deposits99* indicate that bank branches are able to extract greater personal savings from each dollar of income. Given my main argument that distance to a lender affects the effectiveness of bank-customer relations, I reduce the impact of financial development on lending with increasing distance of the branch.

Efficiency00 is the average X-efficiency of the banks present in the market in the 1997-2000 period, calculated using the alternative profit efficiency approach and Fourier-flexible functional form. I prefer the alternative profit efficiency approach over the standard profit efficiency because it provides a way of controlling for unmeasured differences in output quality and market power---which is what information-intensive lending is about. Because the steps involved in calculating the X-efficiency are complicated, I omit that discussion here but refer the reader to Berger and Mester (1997).

Herfindahl00 is the deposit-market Herfindahl index, where the local market is defined as the area within 20 miles of a census tract centroid. Using a 5- or 10-mile radius does not affect the results.

4.4 Sample Selection Issues

The last component of $\boldsymbol{X_1}$ is $\hat{\lambda}$, the Mill's Ratio estimated from the following probit model. It takes into account the effect of the 45 census tracts omitted from the low-income sample due to the lack of spread data. The regression is once again run across the low-income census tracts.

$$\Phi^{-1}\left(\Pr(Event)\right) = \mathbf{X_e}\boldsymbol{\beta}_\mathbf{e} + \varepsilon_{PR} \tag{3}$$

where the *Event* is having the census-tract-level spread reported as zero (all individual spreads less than 3%). $\boldsymbol{X_e}$ includes an intercept term and *Income00, FHA00, Herfindahl00, HighSchool00, HomeValue00, MarketProfit99, ShortCommute00, BAccess99,* and *Race00*. The assumption is that spreads may be low in a market if incomes are high, some loans are FHA-insured, the market is not concentrated, the share of uneducated population is low, collateral values are high, the probability of entry (captured by the profitability of the market *MarketProfit99*) is high, and customers have access to distant banking markets. I also include the racial mix of the community and the level of access to banking services in 1999 as control variables. *MarketProfit99* is the profitability of the banking market in and around the census tract as described in Appendix A.

For the sake of brevity, I do not present these results but I calculate the Mill's Ratio as $\hat{\lambda} = \phi(X_e\hat{\beta}_e) \big/ \left(1 - \Phi(X_e\hat{\beta}_e)\right)$ for each census tract.

The descriptive statistics of all the variables used in the analysis are presented in Table 1. A comparison of sample averages indicates that low-income neighborhoods, compared to middle-high income neighborhoods, have more African-American residents, fewer children who live in a two-parent household, and fewer owner-occupied housing units. Financial institutions are more likely to carry these recently originated loans on their balance sheets (*UnsoldLoan00* in Table 1 Panel B and C), which suggests greater borrower opacity. Also note that the branch access variable (*BAccess*) is greater in low-income neighborhoods compared to higher-income census tracts. This is because high-income neighborhoods are mostly in the suburbs and low-income neighborhoods are closer to urban, business districts with many bank branches.

The model in (2) assumes that branch presence has a lagging effect on mortgage availability. However, there may also be a concurrent effect. Therefore, I repeat the analysis, replacing *BAccess99* and *Deposits99* with *BAccess04* and *Deposits04* and initially treating the latter two as exogenous. However, given the strong potential for endogeneity, I also estimate the following system with GMM:

$$
\begin{aligned}
\textit{Originate04} &= f\,(\textit{BAccess04, Spread04, Deposits04, Retired00, CreditProblem00, OOHousing00}, \\
&\qquad \boldsymbol{X_2}) + \varepsilon_O \\
\textit{Spread04} &= f\,(\textit{BAccess04, Originate04, Deposits04, LenderCost00, CreditProblem00}, \\
&\qquad \textit{OOHousing00}, \boldsymbol{X_2}) + \varepsilon_S \qquad\qquad (4) \\
\textit{BAccess04} &= f\,(\textit{BAccess99, Deposits99, Retired00, LenderCost00, OOHousing00}, \boldsymbol{X_3}) + \varepsilon_B \\
\textit{Deposits04} &= f\,(\textit{BAccess99, Deposits99, Retired00, LenderCost00, CreditProblem00}, \boldsymbol{X_3}) + \varepsilon_D
\end{aligned}
$$

Table 2 lists all the variables on the right-hand side of the equations in (4) to make it easier to track which variable belongs to which equation. *BAccess04* is identified by the assumption that the bank's entry decision will not be affected by the mortgage borrowers' observable past credit histories (*CreditProblem00*). Note that the sample includes only low-income neighborhoods. In this environment, credit information either indicates a low credit quality or is unavailable for a particular class of borrowers (Stegman, Quercia, and Lobenhofer, 2001). Therefore, I assume that while credit quality may affect the branching decision between high-income and low-income neighborhoods, it will have no impact conditional on all the potential markets being low-income. For identification, I also assume that *CreditProblem00* is negatively correlated with *Deposits04* because in markets where credit histories are poor, savings rate will also be low---otherwise, people would have paid their debt. *Deposits04* is identified with the assumption that the share of owner-occupied housing units in the neighborhood's total housing stock, *OOHousing00,* is uncorrelated with the savings rate (deposits per dollar of income) of the population in 2004.

Once again, I have a long list of control variables. $\boldsymbol{X_2}$ contains all the variables in $\boldsymbol{X_1}$ except *Deposits99*, *CreditProblem00*, and *OOHousing00*, which are still in the analysis but I show them explicitly in (4) instead of hiding them in $\boldsymbol{X_1}$.

$\boldsymbol{X_3}$ contains all the variables in $\boldsymbol{X_2}$ and a measure of market profitability in 1994 and 1999 (*MarketProfit94* and *MarketProfit99*), which may affect banks' entry decisions.[11] I also include the 1990 values of some of my control variables, *HighSchool90* and *Manufacturing90* under the assumption that the entry decisions are based on long-term trends in the population's job opportunities and the employer mix of the area.

5. Results

Table 3 presents the results from the system in (2). At the bottom of the Table (and all other Tables that will follow), I present the R-square measure proposed by Windmeijer (1995), which is the squared-correlation of the observed and predicted dependent variables. In low-income areas, greater access to bank branches is associated with improved access to mortgages in 2004, as measured by higher quantity (*Originate04*) and lower cost (*Spread04*). One standard-deviation-increase in *BAccess99* is associated with a \$499 increase in mortgage originations per household (the sample mean is \$3,000) and a 30–basis-point decline in spreads. An interesting observation is that the number of institutions in the market (*Institutions00,* which includes both bank and non-bank lenders) is irrelevant. This finding suggests that it does not matter whether all branches belong to a single bank or whether each branch is an independent entity. I also find that the presence of other types of lenders does not have any statistically significant impact on mortgage availability (*BankShare00*).

Glancing over the other variables, I do not find any sign of discrimination based on race. In fact, the opposite is true. Mortgage availability seems to be higher in minority neighborhoods (*Race00*); originations increase and spreads decline with increasing African-American population. However, if I remove the educational attainment (*HighSchool00*) and the impact of single parents (*ChildInFamily00*) from the analysis (results not shown), the positive correlation between *Race00* and originations

[11] The initial value of market profitability is from 1994 because that is the earliest year for which the FDIC reports the Summary of Deposits on its website.

disappears. More importantly, the correlation between *Race00* and spreads becomes *significantly positive*. This finding suggests that without proper controls, one could mistakenly observe that minorities are discriminated against in the mortgage market. The average efficiency of the institutions in the market seems to be positively correlated with originations and negatively correlated with spreads but these effects are not statistically significant. As I mentioned earlier, the ability to access distant banking markets (*ShortCommute00* measures lack of access) may have either a positive effect on mortgage availability due to disappearing monopoly rents in the local market, or a negative effect because of weakening relationships under increased competition. My analysis suggests that the latter effect dominates. As people work closer to home (high *ShortCommute00*), mortgage originations tend to increase and spreads tend to decrease. This is a strong support for the relationship-lending theory. Finally, I find that tract residents' educational attainment has a significant impact on mortgage availability. I interpret education attainment as a proxy for the accessibility of gainful and stable jobs. As educational attainment deteriorates (high *HighSchool00*) across census tracts, mortgage originations drop and spreads increase.

Table 4 repeats the analysis but this time using the contemporaneous branch access variable *BAccess04*. The conclusions are the same but the economic effect of branch access is stronger. This time, one standard deviation increase in *BAccess04* is associated with a $637 increase in originations per household and a 39 basis point drop in spreads. However, this analysis does not account for the endogeneity of the branching decision. Therefore, in Table 5, I endogenize branch access and financial development as measured by *Deposits04* using the system in (4). The results are very strong. A one–standard-deviation increase in branch access is associated with a $868 increase in originations per household and a 56–basis-point drop in spreads. The increase of $868 corresponds to a 29% increase in originations relative to sample mean of *Originate04*. Also note from Table 1 that the average spread over Treasuries in the sample is 1.9%. The decline in spreads by 56 bp is a 29% drop from the sample mean. These findings support the first and second hypotheses of the paper.

5.1 Robustness Checks

I have so far tested the implications of the relationship lending theory by investigating whether those who are most likely to benefit from relationships actually do so. An alternative way of looking at this problem is to investigate which group of individuals would benefit *less* from relationships in the mortgage market. Anticipating that information-driven problems will be less severe in high-income areas, it is safe to assume that mortgage lending in these areas will depend mostly on credit scoring, not relationships. If I find that branch presence is important for mortgage availability in both low- and high-income neighborhoods, I can no longer make the argument that what drives the importance of branches is relationship lending. To test this hypothesis, I estimate system (4) in high-income census tracts. The results are in Table 6, Panel A. As expected, access to a bank branch has no significant impact on originations and spreads in high-income areas.

By design, *BAccess* assumes that the impact of a branch on credit availability declines with its distance to the market. If this assumption is correct, disregarding the distance and simply counting the bank branches in and around a neighborhood should weaken the results statistically and economically. Weaker results would indicate that the distance of the lender is important and that relationships do exist.

I will demonstrate the importance of distance in two ways. First, I will ignore the distance of the branch to the census tract centroid. In other words, if I find a branch within ten miles of the centroid, I will increase the access variable by one irrespective of the distance. Second, I will still use the 1/Distance ratio in calculating the access variable but I will decrease the radius of my search to five miles and then repeat the analysis after increasing it to 20 miles. If the 10-mile radius is the correct distance to consider for viable relationships, adding more distant branches to the calculation should increase the noise in the measure and weaken my results.

The results are in Table 6 Panels B and C. Panel B shows that, as expected, the economic effect of *BAccess04* weakens when I increase the radius to 20 miles. The effect also weakens when I decrease the radius to five miles, but the decline is much smaller. Panel C shows that if I ignore the distance completely and just add up the branches, the conclusions may be misleading. According to Panel C (using the 10-mile radius results), an increase in the local number of branches from zero to one, 10 miles from the centroid, increases originations per household by $9,165. However, if we account for the distance of the branch, the actual increase in the branch-access variable is 0.1 (1/10). According to Table 5, the effect of such an increase is a more modest $4,177. Conversely, if the new branch is only 1 mile from the centroid, Panel C of Table 6 would still forecast an increase in originations by $9,165. However, at such a close distance, Table 5 suggests that the increase is $30,380. But which one is the correct prediction? The one that includes the distance or the one that ignores it? To answer this question, I estimate the following system:

$$
\begin{aligned}
&\textit{Originate04} &&= f\,(\textit{BAccess04, BAccess04id, Spread04, Deposits04, Retired00, CreditProblem00,}\\
& &&\quad \textit{OOHousing00, } \boldsymbol{X_2}) + \varepsilon_{O}\\
&\textit{Spread04} &&= f\,(\textit{BAccess04, BAccess04id, Originate04, Deposits04, LenderCost00,}\\
& &&\quad \textit{CreditProblem00, OOHousing00, } \boldsymbol{X_2}) + \varepsilon_{S} \qquad\qquad (5)\\
&\textit{BAccess04} &&= f\,(\textit{BAccess99, Deposits99, Retired00, LenderCost00, OOHousing00, } \boldsymbol{X_3}) + \varepsilon_{B}\\
&\textit{BAccess04id} &&= f\,(\textit{BAccess99id, Deposits99, Retired00, LenderCost00, OOHousing00, } \boldsymbol{X_3}) + \varepsilon_{Bd}\\
&\textit{Deposits04} &&= f\,(\textit{BAccess99, Deposits99, Retired00, LenderCost00, CreditProblem00, } \boldsymbol{X_3}) + \varepsilon_{D}
\end{aligned}
$$

where *BAccess04id* is the natural log of one plus the number of bank branches within 10 miles of the census tract centroid, irrespective of their distance. I include both *BAccess04id* and *BAccess04* in the origination and spread equations to find out which one will survive. The results are in Table 6, Panel D. *BAccess04* is the significant variable. These findings suggest that ignoring the distance between a neighborhood and lenders may lead to overestimating the impact of a distant branch and underestimating the impact of a branch nearby.

To investigate whether my results are impacted by outliers, I repeat the analysis in two different ways. First, I re-run regression (4) after winsorizing *BAccess04*, *Originate04*, and *Spread04* at the 1% and 99% levels to reduce the impact of potential outliers.[12] Second, I redefine *BAccess04* as in (1'). The results are in Table 6 Panel E. While the results are somewhat weaker, the main conclusions stand.

[12] Any observation that is below the 1st percentile or above the 99th percentile of the sample is reset to the 1st percentile and 99th percentile values.

Next, I run my final robustness check. As I mentioned earlier, I may be making a mistake in calculating *BAccess* in rural areas with large census tracts and small population centers. To reduce the impact of mismeasured *BAccess*, I rank all my low-income census tracts by declining population density and remove the lowest quartile from the sample. The results from this smaller sample are in Table 6 Panel F. There is no material difference between these results and the results in Table 5.

6. Conclusion

Relationship-lending literature has shown that bank-borrower interactions that reveal private information about the borrower to the lender are important in improving small businesses' access to credit. In this paper, I apply the same theory to lending in low-income neighborhoods. I hypothesize that the presence of a bank branch in a low-income community will improve the access to mortgage loans in the area by reducing the distance-related frictions in the information gathering process. Thus, I investigate the connection between a branch access measure that I develop and mortgage originations and spreads. I find a strong positive relationship between branch access and originations and a strong negative relationship between branch access and spreads.

While several problems associated with HMDA data necessitates caution while interpreting the results, my findings still suggest that the usefulness of bank-borrower relationships is not limited to the small business lending market, where the focus of the literature has been.

References

Adams, D. and J.D. von Pischke, 1992, "Microenterprise credit programs: Déjà vu", *World Development* 20, p. 1463-1470.

Agarwal, S., and R. Hauswald, 2006, "Distance and information asymmetries in lending decisions", unpublished manuscript, April 2006.

Akhavein, J., S. Frame, and L.J. White, 2005, "The diffusion of financial innovations: An examination of the adoption of small business credit scoring by large banking organizations", *Journal of Business* 78, p. 577-596.

Amel, D.F., and N. Liang, 1997, "Determinants of entry and profits in local banking markets", *Review of Industrial Organization* 12, p. 59-78.

Anderson, S., J. Baland, and K.O. Moene, 2003, "Sustainability and organizational design in informal groups: Someevidence from Kenyan Roscas", University of Oslo, Department of Economics, Memorandum No. 17.

Armendáriz de Aghion, B. and J. Morduch, 2004, "The economics of microfinance", MIT Press, Cambridge: MA.

Avery, R.B., P.E. Beeson, and M.S. Sniderman, 1999, "Neighborhood Information and Home Mortgage Lending", *Journal of Urban Economics* 45, p. 287-310.

Beck, T., A. Demirguc-Kunt, and R. Levine, 2004, "Finance, Inequality, and Poverty: Cross-Country Evidence", *NBER Working Paper #: 10979*.

Beegle, K., R.H. Dehejia, and R. Gatti, 2003, "Child labor, income shocks, and access to credit", World Bank Policy Research Working Paper No. 3075.

Bencivenga, V.R. and B.D. Smith, 1991, "Financial intermediation and endogenous growth", *Review of Economic Studies* 58, p. 195-209.

Berger, A.N., and L.J. Mester, 1997, "Inside the black box: What explains differences in the efficiencies of financial institutions?", *Journal of Banking and Finance* 21, p. 895-947.

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

Berger, A. N., and G. F. Udell, 1995, "Relationship lending and lines of credit in small business finance", *Journal of Business* 68, p. 351-381.

Berger, A. N., and G. F. Udell, 1996. Universal banking and the future of small business lending, in: Saunders, A., Walter, I. (Eds.), Universal Banking: Financial System Design Reconsidered, Irwin, Burr Ridge: Illinois, 559-627.

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

Besley, T., and S. Coate, 1995, "Group lending, repayment incentives, and social collateral", *Journal of Development Economics* 46, p. 1-18.

Besley, T., S. Coate, and G. Loury, 1993, "The economics of rotating savings and credit associations", *American Economic Review* 83, p. 792-810.

Brevoort, K.P., and T.H. Hannan, 2004, "Commercial lending and distance: Evidence from Community Reinvestment Act data", Federal Reserve Board Finance and Economics Discussion Series No. 2004-24.

Chomsisengphet, S., and A. Pennington-Cross, 2006, "The evolution of the subprime mortgage market", *Federal Reserve Bank of St. Louis Review*, January/February 2006, p. 31-56.

Christen, R.P., E. Rhyne, R. Vogel, and C. McKean, 1995, "Maximizing the outreach of microenterprise finance: The emerging lessons of successful programs", USAID Program and Operations Assessment Report No. 10, Washington, D.C.

Cole, R.A., L.G. Goldberg and L.J. White, 2004, "Cookie-Cutter versus Character: The Micro Structure of Small Business Lending by Large and Small Banks", *Journal of Financial and Quantitative Analysis* 39, 227-251.

Coleman, B.E., 1999, "The impact of group lending in Northeast Thailand", *Journal of Development Economics* 60, p. 105-141.

Coleman, B.E., 2002, “Microfinance in Northeast Thailand: Who benefits and how much?”, Asian Development Bank ERD Working Paper No. 9.

Conlin, M., 1999, “Peer group micro-lending programs in Canada and the United States”, *Journal of Development Economics* 60, p. 249-269.

Degryse H. and S. Ongena, 2005, “Distance, lending relationships, and competition”, *Journal of Finance* 60, p. 231-266.

Demetriades, P.O., and K.A. Hussein , 1996, “Does financial development cause economic growth? Time-series evidence from 16 countries”, *Journal of Development Economics* 51, p. 387-411.

DeYoung, R., D. Glennon, and P. Nigro, 2006, “Borrower-lender distance, credit scoring, and the performance of small business loans”, FDIC Center for Financial Research Working Paper No. 2006-4.

Frame, W. S., A. Srinivasan, L. Woosley, 2001, “The effect of credit scoring on small business lending”, *Journal of Money, Credit, and Banking* 33, 813-825.

Galor, O. and J. Zeira, 1993, “Income distributions and macroeconomics”, *Review of Economic Studies* 60, p. 35-52.

Ghatak, M., 1999, “Group lending, local information and peer selection”, *Journal of Development Economics* 60, 27-50.

Ghosh, P., and D. Ray, 1996, “Cooperation in community interaction without information flows”, *Review of Economic Studies* 63, p. 491-519.

Goetz, A.M., and R.S. Gupta, 1996, “Who takes the credit? Gender, power, and control over loan use in rural credit programs in Bangladesh”, *World Development* 24, p. 45-63.

Guiso, L., P. Sapienza, and L. Zingales, 2002, “Does local financial development matter?”, *NBER Working Paper #: 8923*.

Hauswald, R. and R. Marquez, 2006, “Competition and strategic information acquisition in credit markets”, *Review of Financial Studies* 19, p. 967-1000.

Honohan, P., 2004a, “Financial sector policy and the poor: Selected findings and issues”, World Bank Working Paper No. 43.

Honohan, P., 2004b, “Financial development, growth and poverty: How close are the links?”, World Bank Policy Research Working Paper No. 3203.

Jain, P., 1996, “Managing credit for rural poor: Lessons from the Grameen Bank”, *World Development* 24, p. 78-89.

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

Jayaratne, J. and J. Wolken, 1999, “How important are small banks to small business lending? New evidence from a survey of small firms,” *Journal of Banking and Finance* 23, p. 427–458.

King, R.G. and R. Levine, 1993, “Finance and Growth: Schumpeter Might Be Right”, *Quarterly Journal of Economics* 108, p. 717-738.

Lederer, P., and A.P. Hurter, 1986, “Competition of firms: Discriminatory pricing and location”, *Econometrica* 54, p. 623-640.

Levonian, M. E., 1997, “Changes in small business lending in the West”, Federal Reserve Bank of San Francisco Economic Letter no. 97-02.

Matin, I., D. Hulme, and S. Rutherford, 2002, “Finance for the poor: From microcredit to microfinancial services”, *Journal of International Development* 14, p. 273-294.

Mester, L., L.I. Nakamura, and M. Renault, 1998, “Checking accounts and bank monitoring”, Federal Reserve Bank of Philadelphia Working Paper No. 98-25 .

MkNelly, B., and C. Dunford, 1998, “Impact of credit with education on mothers and their young children’s nutrition: Lower Pra Rural Bank credit with education program in Ghana”, Freedom from Hunger Research Paper No. 4.

MkNelly, B., and C. Dunford, 1999, “Impact of credit with education on mothers and their young children’s nutrition: CRECER credit with education program in Bolivia”, Freedom from Hunger Research Paper No. 5.

Morduch, J., 1999, “The microfinance promise”, *Journal of Economic Literature* 37, p. 1569-1614.

Peek, J. and E.S. Rosengren, 1998, “Bank consolidation and small business lending: It is not just bank size that matters,” *Journal of Banking and Finance* 22, p. 799–819

Petersen, M.A., and R.G. Rajan, 1994, “The benefits of lending relationships: Evidence from small business data”, *Journal of Finance* 49, p. 3-37.

Petersen, M.A., and R.G. Rajan, 1995, “The effect of credit market competition on lending relationships”, *Quarterly Journal of Economics* 110, p. 407-442.

Petersen M.A., and R. Rajan, 2002, “Does distance still matter? The information revolution in small business lending”, *Journal of Finance* 57, p. 2533-2570.

Pulley, R., 1989, “Making the poor creditworthy: A case study of the integrated rural development program in India”, World Bank Discussion Paper No. 48.

Rajan, R.G., 1992, “Insiders and outsiders: The choice between informed and arm's-length debt”, Journal of Finance 47, p. 1367-1400.

Sharpe, S.A., 1990, “Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships”, *Journal of Finance* 45, p. 1069-1087.

Sinnott, R.W., 1984, "Virtues of the Haversine", Sky and Telescope 68, no. 2, p. 159.

Stegman, M.A., R.G. Quercia, and J.S. Lobenhofer, 2001, “Underwriting: Getting to ‘yes’ for more low-income applicants”, 2001 Conference on Housing Opportunity: Will Technology Expand Housing Opportunity?

Stein, J.C., 2002, “Information production and capital allocation: Decentralized vs. hierarchical firms”, *Journal of Finance* 57, p. 1891-1921.

Stiglitz, J., and A. Weiss, 1981, “Credit rationing in markets with imperfect information”, *American Economic Review* 71, p. 393-410.

Stiglitz, J., 1990, “Peer monitoring and credit markets”, *World Bank Economic Review* 4, p. 351-366.

Strahan, P.E. and J.P. Weston, 1998, “Small business lending and the changing structure of the banking industry,” *Journal of Banking and Finance* 22, p. 821–845.

Varian, H., 1990, “Monitoring agents with other agents”, *Journal of Institutional and Theoretical Economics* 146, p. 153-174.

von Thadden, E.-L., 2002, “Asymmetric information, bank lending and implicit contracts: The winner’s curse”, Unpublished working paper, Universite de Lausanne.

Windmeijer, F., 1995, “A Note on R^2 in the instrumental variables model”, Journal of Quantitative Economics 11, p. 257-261.

Wolfensohn, J.D., and F. Bourguignon, 2004, “Development and poverty reduction: Looking back, looking ahead”, *The World Bank*, mimeo.

Wydick, W.B., 1999, “Can social cohesion be harnessed to repair market failures? Evidence from group lending in Guatemala”, *Economic Journal* 109, p. 463-475.

Table 1: Descriptive Statistics
Panel A. All Observations (N=2023)

	Mean	Std Dev	Median	Minimum	Maximum
Endogenous Variables					
BAccess04	0.0139	0.0240	0.0083	0.0000	0.5800
Deposits04	2.6796	1.6903	2.5216	0.0000	14.7325
Originate04	4.2723	4.4293	3.1008	0.0000	56.5091
Spread04	1.5795	0.9058	1.4541	0.0000	6.3700
Demographic Factors					
ChildInFamily00	0.4816	0.1301	0.5242	0.0253	0.6931
HighSchool00	0.1718	0.0873	0.1587	0.0000	0.5653
HighSchool90	0.2278	0.0997	0.2191	0.0000	0.6317
Income00	3.4030	0.7437	3.5633	-6.9078	5.2983
Income90	3.2784	0.6783	3.4457	-1.9661	5.0349
Jobs00	0.3785	0.0564	0.3929	0.0494	0.5692
Manufacturing00	0.1856	0.0689	0.1807	0.0000	0.4502
Manufacturing90	0.2227	0.0747	0.2205	0.0115	0.4869
Population00	0.1271	1.7112	0.5946	-4.4896	3.0609
Population90	0.1123	1.7677	0.5757	-4.1879	3.6009
Race00	0.1329	0.2015	0.0300	0.0000	0.6904
Retired00	0.1737	0.0500	0.1741	0.0000	0.3665
RuralPop00	0.2259	0.3653	0.0000	0.0000	1.0000
ShortCommute00	0.5434	0.0647	0.5552	0.0648	0.6931
Mortgage-Market-Specific Factors					
Applications00	0.0801	0.0572	0.0721	0.0000	1.6667
BankShare00	0.6691	0.1394	0.6824	0.0000	1.0000
BorrowerIncome00	3.8646	0.3287	3.8131	2.7726	5.7666
CoApplicant00	0.5002	0.1637	0.5255	0.0000	1.0000
Conventional00	0.0007	0.0036	0.0002	0.0000	0.0953
CreditProblem00	0.0748	0.0608	0.0612	0.0000	0.5000
FHA00	0.1414	0.1235	0.1059	0.0000	0.7082
HomeValue00	3.7205	1.1558	3.9228	-4.4804	9.2904
HomeValue90	3.4967	1.1286	3.6746	-4.1073	9.2434
Institutions00	3.5773	0.5707	3.6109	0.0000	5.4638
OOHousing00	0.7190	0.1973	0.7672	0.0039	1.0714
Originate00	3.2923	3.1260	2.4375	0.0000	37.2477
UnsoldLoan00	0.6216	0.1477	0.6190	0.0000	1.0000
Banking-Market-Specific Factors					
BAccess99	0.0207	0.0344	0.0095	0.0000	0.5401
Deposits99	2.6907	1.7612	2.5156	0.0000	14.8561
Efficiency00	0.7160	0.0636	0.7149	0.4832	0.9242
Herfindahl00	0.1371	0.0396	0.1392	0.0599	0.3909
LenderCost99	0.0503	0.0121	0.0459	0.0345	0.0865
MarketProfit94	0.0350	0.0103	0.0354	0.0087	0.0663
MarketProfit99	0.0466	0.0138	0.0466	0.0080	0.0934

Panel B. Low-Income Neighborhoods excluding zero-spread census tracts (N=1244)

	Mean	Std Dev	Median	Minimum	Maximum
Endogenous Variables					
BAccess04	0.0136	0.0198	0.0084	0.0000	0.2809
Deposits04	2.9985	1.7564	2.8644	0.0232	14.7325
Originate04 ($000)	2.9916	3.0859	2.1638	0.0000	35.1524
Spread04 (%)	1.9058	0.8970	1.7864	0.0113	6.3700
Demographic Factors					
ChildInFamily00	0.4353	0.1288	0.4703	0.0253	0.6748
HighSchool00	0.2048	0.0829	0.1944	0.0267	0.5653
HighSchool90	0.2627	0.0931	0.2577	0.0251	0.5919
Income00	3.0986	0.7674	3.3072	-6.9078	3.7124
Income90	2.9930	0.6647	3.1783	-1.9661	4.2408
Jobs00	0.3640	0.0568	0.3734	0.1220	0.5692
Manufacturing00	0.1875	0.0672	0.1844	0.0150	0.4502
Manufacturing90	0.2216	0.0745	0.2191	0.0351	0.4869
Population00	0.3991	1.6550	0.9073	-4.2134	3.0609
Population90	0.4122	1.7121	0.9299	-4.1879	3.6009
Race00	0.1758	0.2221	0.0624	0.0003	0.6904
Retired00	0.1688	0.0480	0.1710	0.0000	0.3606
RuralPop00	0.1645	0.3110	0.0000	0.0000	1.0000
ShortCommute00	0.5471	0.0625	0.5595	0.0648	0.6686
Mortgage-Market-Specific Factors					
Applications00	0.0806	0.0386	0.0741	0.0090	0.6265
BankShare00	0.6417	0.1398	0.6502	0.0000	1.0000
BorrowerIncome00	3.7501	0.2735	3.7035	2.7726	5.2661
CoApplicant00	0.4494	0.1576	0.4700	0.0000	0.8065
Conventional00	0.0007	0.0032	0.0003	0.0000	0.0953
CreditProblem00	0.0898	0.0629	0.0769	0.0000	0.5000
FHA00	0.1567	0.1270	0.1283	0.0000	0.6213
HomeValue00	3.3059	1.2173	3.6714	-4.4804	6.1193
HomeValue90	3.0600	1.1422	3.3806	-4.1073	5.3583
Institutions00	3.6144	0.5612	3.6109	1.0986	5.4638
OOHousing00	0.6584	0.1881	0.6883	0.0348	0.9935
Originate00	2.2961	2.0691	1.7671	0.0000	20.8177
UnsoldLoan00	0.6541	0.1416	0.6667	0.2857	1.0000
Banking-Market-Specific Factors					
BAccess99	0.0211	0.0351	0.0100	0.0000	0.5401
Deposits99	3.0007	1.8209	2.8488	0.0240	14.8561
Efficiency00	0.7195	0.0653	0.7226	0.4832	0.9242
Herfindahl00	0.1383	0.0412	0.1407	0.0614	0.3909
LenderCost99	0.0501	0.0120	0.0459	0.0345	0.0865
MarketProfit94	0.0355	0.0105	0.0363	0.0090	0.0663
MarketProfit99	0.0471	0.0136	0.0466	0.0080	0.0934

Panel C. Middle-High Income Neighborhoods excluding zero-spread census tracts (N=710)

	Mean	Std Dev	Median	Minimum	Maximum
Endogenous Variables					
BAccess04	0.0099	0.0089	0.0078	0.0001	0.0590
Deposits04	1.9492	1.1132	1.8862	0.0434	6.5250
Originate04 ($000)	6.5981	5.3507	5.0536	0.5135	56.5091
Spread04 (%)	1.1605	0.5699	1.1199	0.0119	4.9182
Demographic Factors					
ChildInFamily00	0.5699	0.0541	0.5758	0.2621	0.6830
HighSchool00	0.1116	0.0469	0.1143	0.0040	0.2719
HighSchool90	0.1634	0.0629	0.1688	0.0142	0.3461
Income00	3.9405	0.1996	3.8831	3.7126	5.0808
Income90	3.7902	0.2567	3.7530	0.4935	4.9558
Jobs00	0.4086	0.0281	0.4088	0.1604	0.4934
Manufacturing00	0.1873	0.0695	0.1794	0.0327	0.3883
Manufacturing90	0.2300	0.0722	0.2277	0.0509	0.4290
Population00	-0.3895	1.6824	-0.1396	-3.9542	2.6326
Population90	-0.4641	1.7236	-0.2929	-3.8991	2.6512
Race00	0.0479	0.1070	0.0105	0.0003	0.6751
Retired00	0.1852	0.0469	0.1809	0.0000	0.3665
RuralPop00	0.3380	0.4222	0.0367	0.0000	1.0000
ShortCommute00	0.5375	0.0628	0.5454	0.3193	0.6666
Mortgage-Market-Specific Factors					
Applications00	0.0742	0.0360	0.0691	0.0113	0.5000
BankShare00	0.7106	0.1161	0.7235	0.2274	0.9506
BorrowerIncome00	4.0524	0.2996	3.9811	3.4741	5.3888
CoApplicant00	0.5939	0.1125	0.6000	0.1579	1.0000
Conventional00	0.0002	0.0003	0.0001	0.0000	0.0077
CreditProblem00	0.0492	0.0409	0.0395	0.0000	0.2464
FHA00	0.1208	0.1133	0.0865	0.0000	0.7082
HomeValue00	4.3170	0.5172	4.2443	2.9764	7.3513
HomeValue90	4.1478	0.5468	4.0652	1.0356	7.4712
Institutions00	3.6238	0.4093	3.6636	1.3863	4.7791
OOHousing00	0.8424	0.1154	0.8693	0.0167	1.0371
Originate00	5.0242	3.6281	4.0999	0.1656	37.2477
UnsoldLoan00	0.5578	0.1272	0.5433	0.2264	1.0000
Banking-Market-Specific Factors					
BAccess99	0.0140	0.0159	0.0079	0.0001	0.1318
Deposits99	1.9651	1.2439	1.7514	0.0350	6.7921
Efficiency00	0.7088	0.0591	0.7065	0.5955	0.8860
Herfindahl00	0.1343	0.0369	0.1359	0.0599	0.2871
LenderCost99	0.0504	0.0123	0.0459	0.0355	0.0865
MarketProfit94	0.0345	0.0099	0.0345	0.0087	0.0611
MarketProfit99	0.0464	0.0142	0.0468	0.0085	0.0848

Table 2: List of Regressors

BAccess04	**Deposits04**	**Originate04**	**Spread04**
Applications00	Applications00	Applications00	Applications00
BankShare00	BankShare00	BankShare00	BankShare00
BorrowerIncome00	BorrowerIncome00	BorrowerIncome00	BorrowerIncome00
ChildInFamily00	ChildInFamily00	ChildInFamily00	ChildInFamily00
CoApplicant00	CoApplicant00	CoApplicant00	CoApplicant00
Conventional00	Conventional00	Conventional00	Conventional00
Efficiency00	Efficiency00	Efficiency00	Efficiency00
FHA00	FHA00	FHA00	FHA00
Herfindahl00	Herfindahl00	Herfindahl00	Herfindahl00
HighSchool00	HighSchool00	HighSchool00	HighSchool00
HomeValue00	HomeValue00	HomeValue00	HomeValue00
HomeValue90	HomeValue90	HomeValue90	HomeValue90
Income00	Income00	Income00	Income00
Income90	Income90	Income90	Income90
Institutions00	Institutions00	Institutions00	Institutions00
Jobs00	Jobs00	Jobs00	Jobs00
$\hat{\lambda}$	$\hat{\lambda}$	$\hat{\lambda}$	$\hat{\lambda}$
Manufacturing00	Manufacturing00	Manufacturing00	Manufacturing00
Originate2000	Originate2000	Originate2000	Originate2000
Population00	Population00	Population00	Population00
Population90	Population90	Population90	Population90
Race00	Race00	Race00	Race00
RuralPop00	RuralPop00	RuralPop00	RuralPop00
ShortCommute00	ShortCommute00	ShortCommute00	ShortCommute00
UnsoldLoan00	UnsoldLoan00	UnsoldLoan00	UnsoldLoan00
	CreditProblem00	CreditProblem00	CreditProblem00
OOHousing00		OOHousing00	OOHousing00
LenderCost99	LenderCost99		LenderCost99
Retired00	Retired00	Retired00	
MarketProfit99	MarketProfit99		
HighSchool90	HighSchool90		
Manufacturing90	Manufacturing90		
MarketProfit94	MarketProfit94		
BAccess99	BAccess99		
Deposits99	Deposits99		
		Spread04	
			Originate04
		BAccess04	BAccess04
		Deposits04	Deposits04

Table 3: The Effect of Branch Presence on Mortgage Availability in Low - Moderate Income Neighborhoods – GMM with Exogenous Branch Presence in 1999

This table shows the effect of bank branch presence in 1999, *BAccess99*, on mortgage availability by estimating the following system with GMM:

Originate04 = f (*BAccess99, Spread04, Retired00,* $\boldsymbol{X_1}$) + ε_O

Spread04 = f (*BAccess99, Originate04, LenderCost00,* $\boldsymbol{X_1}$) + ε_S

R-square is the squared-correlation of the observed endogenous variable with its predicted value.

t-statistics are in parenthesis. Some of the variables in $\boldsymbol{X_1}$ are omitted from the Table for expositional reasons.

(Table on next page)

	Originate2004	Spread2004
BAccess99	25.186	-15.404
	(4.01) ***	(-3.21) ***
Deposits99	-0.105	0.066
	(-0.90)	(0.96)
Race00	1.738	-1.081
	(1.94) *	(-1.72) *
ShortCommute00	4.460	-2.818
	(2.91) ***	(-2.79) ***
BankShare00	0.433	-0.351
	(0.80)	(-1.09)
Income90	-0.627	0.377
	(-1.45)	(1.41)
Income00	-0.289	0.202
	(-0.48)	(0.56)
HighSchool00	-6.660	4.314
	(-3.16) ***	(3.17) ***
Institutions00	0.040	-0.021
	(0.15)	(-0.13)
HomeValue00	-0.459	0.250
	(-1.03)	(0.94)
UnsoldLoan00	1.268	-0.650
	(1.89) *	(-1.44)
ChildInFamily00	1.036	-1.093
	(0.51)	(-0.92)
Jobs00	4.859	-2.969
	(2.20) **	(-2.29) **
Manufacturing00	-2.391	1.380
	(-2.17) **	(2.25) **
CreditProblem00	0.206	-0.093
	(0.22)	(-0.16)
LenderCost00		-1.788
		(-1.07)
Retired00	-0.303	
	(-0.45)	
Efficiency00	1.128	-0.715
	(1.37)	(-1.47)
Herfindahl00	-1.986	1.209
	(-1.30)	(1.26)
RuralPop00	-0.376	0.242
	(-1.18)	(1.21)
$\hat{\lambda}$	3.041	-1.930
	(3.83) ***	(-3.53) ***
R-square	0.72	0.26

(***), (**), and (*) denote significant at 1%, 5%, and 10% level, respectively.

Table 4: The Effect of Branch Presence on Mortgage Availability in Low - Moderate Income Neighborhoods – GMM with Exogenous Branch Presence in 2004

This table shows the effect of bank branch presence in 2004, *BAccess04*, on mortgage availability by estimating the following system with GMM:

Originate04 $= f\,(BAccess04,\ Spread04,\ Retired00,\ \boldsymbol{X_1}) + \varepsilon_O$

Spread04 $= f\,(BAccess04,\ Originate04,\ LenderCost00,\ \boldsymbol{X_1}) + \varepsilon_S$

R-square is the squared-correlation of the observed endogenous variable with its predicted value.

t-statistics are in parenthesis. Some of the variables in $\boldsymbol{X_1}$ are omitted from the Table for expositional reasons.

(Table on next page)

	Originate2004	Spread2004
BAccess04	32.152	-19.863
	(3.56) ***	(-2.56) **
Deposits04	-0.171	0.113
	(-1.56)	(1.95) *
Race00	0.844	-0.485
	(1.01)	(-0.81)
ShortCommute00	2.437	-1.702
	(1.88) *	(-1.93) *
BankShare00	0.673	-0.513
	(1.17)	(-1.60)
Income90	-0.547	0.359
	(-1.31)	(1.39)
Income00	-0.456	0.265
	(-0.78)	(0.78)
HighSchool00	-3.952	2.843
	(-2.41) **	(2.44) **
Institutions00	0.126	-0.062
	(0.46)	(-0.38)
HomeValue00	-0.320	0.188
	(-0.78)	(0.76)
UnsoldLoan00	1.615	-0.697
	(2.50) **	(-1.54)
ChildInFamily00	1.097	-1.179
	(0.55)	(-1.03)
Jobs00	4.309	-2.922
	(1.97) *	(-2.40) **
Manufacturing00	-2.026	1.263
	(-1.83) *	(2.11) **
CreditProblem00	-0.190	0.128
	(-0.21)	(0.23)
LenderCost00		-2.354
		(-1.16)
Retired00	-0.956	
	(-1.11)	
Efficiency00	0.892	-0.651
	(1.12)	(-1.35)
Herfindahl00	-0.294	0.202
	(-0.22)	(0.24)
RuralPop00	-0.324	0.183
	(-1.07)	(0.95)
$\hat{\lambda}$	1.825	-1.269
	(2.93) ***	(-2.64) ***
R-square	0.73	0.28

(***), (**), and (*) denote significant at 1%, 5%, and 10% level, respectively.

Table 5: The Effect of Branch Presence on Mortgage Availability in Low - Moderate Income Neighborhoods – GMM with Endogenous Branch Presence

This table shows the effect of endogenous bank branch presence, *BAccess04*, and endogenous savings rate in 2004, *Deposits04*, on mortgage availability by estimating the following system with GMM:

Originate04 = *f* (*BAccess04, Spread04, Deposits04, Retired00, CreditProblem00, OOHousing00,* $\boldsymbol{X_2}$) + ε_O

Spread04 = *f* (*BAccess04, Originate04, Deposits04, LenderCost00, CreditProblem00, OOHousing00,* $\boldsymbol{X_2}$) + ε_S

BAccess04 = *f* (*BAccess99, Deposits99, Retired00, LenderCost00, OOHousing00,* $\boldsymbol{X_3}$) + ε_B

Deposits04 = *f* (*BAccess99, Deposits99, Retired00, LenderCost00, CreditProblem00,* $\boldsymbol{X_3}$) + ε_D

R-square is the squared-correlation of the observed endogenous variable with its predicted value.

t-statistics are in parenthesis. Some of the variables in $\boldsymbol{X_1}$ are omitted from the Table for expositional reasons.

(Table on next page)

	Originate2004	Spread2004
BAccess04	43.830	-28.185
	(3.99) ***	(-2.69) ***
Deposits04	-0.111	0.066
	(-0.86)	(0.81)
Race00	1.490	-0.933
	(1.74) *	(-1.31)
ShortCommute00	3.232	-2.323
	(2.53) ***	(-2.35) ***
BankShare00	0.550	-0.431
	(0.98)	(-1.25)
Income90	-0.551	0.373
	(-1.36)	(1.36)
Income00	-0.234	0.163
	(-0.39)	(0.43)
HighSchool00	-5.061	3.755
	(-2.81) ***	(2.66) ***
Institutions00	0.064	-0.049
	(0.24)	(-0.27)
HomeValue00	-0.506	0.295
	(-1.19)	(1.07)
UnsoldLoan00	1.417	-0.768
	(2.16) **	(-1.49)
ChildInFamily00	0.532	-0.847
	(0.27)	(-0.69)
Jobs00	4.915	-3.375
	(2.34) ***	(-2.55) ***
Manufacturing00	-2.203	1.424
	(-2.08) ***	(2.26) ***
CreditProblem00	0.218	-0.111
	(0.25)	(-0.19)
LenderCost00		-2.145
		(-1.02)
Retired00	-0.564	
	(-0.61)	
Efficiency00	1.058	-0.760
	(1.33)	(-1.49)
Herfindahl00	-0.705	0.580
	(-0.53)	(0.66)
RuralPop00	-0.230	0.145
	(-0.75)	(0.70)
$\hat{\lambda}$	2.595	-1.801
	(3.69) ***	(-2.91) ***
R-square	0.75	0.24

(***), (**), and (*) denote significant at 1%, 5%, and 10% level, respectively.

Table 6: Robustness Checks

Panel A. The Effect of Branch Presence on Mortgage Availability in Higher-Income Neighborhoods

	Originate2004	Spread2004
BAccess04	11.887	-1.046
	(0.12)	(-0.05)
R-square	0.68	0.20

Panel B. The Effect of Alternative Radii in Low - Moderate Income Neighborhoods

	5 miles		20 miles	
	Originate2004	Spread2004	Originate2004	Spread2004
BAccess04	42.634	-24.214	21.893	-11.611
	(4.49) ***	(-3.66) ***	(4.38) ***	(-3.60) ***
R-square	0.71	0.26	0.70	0.28

Panel C. The Effect of Ignoring Branch Proximity on the Estimated Effect of Branch Presence on Mortgage Availability in Low - Moderate Income Neighborhoods

	5 miles		10 miles	
	Originate2004	Spread2004	Originate2004	Spread2004
BAccess04	11.576	-6.720	13.222	-8.801
	(2.49) **	(-2.16) **	(3.67) ***	(-2.63) ***
R-square	0.72	0.27	0.71	0.24

Panel D. Is Distance Important?

	Originate2004	Spread2004
BAccess04	49.605	-36.676
	(4.30) ***	(-3.42) ***
BAccess04id	-0.452	1.165
	(-0.11)	(0.39)
R-square	0.73	0.20

(***), (**), and (*) denote significant at 1%, 5%, and 10% level, respectively.

Panel E. The Impact of Outliers (Low - Moderate Income Neighborhoods)

	Winsorized Sample		BAccess04 Redefined	
	Originate2004	Spread2004	Originate2004	Spread2004
BAccess04	32.572	-10.743	41.275	-25.369
	(3.82) ***	(-1.64)*	(3.79) ***	(-2.78) ***
R-square	0.78	0.49	0.72	0.25

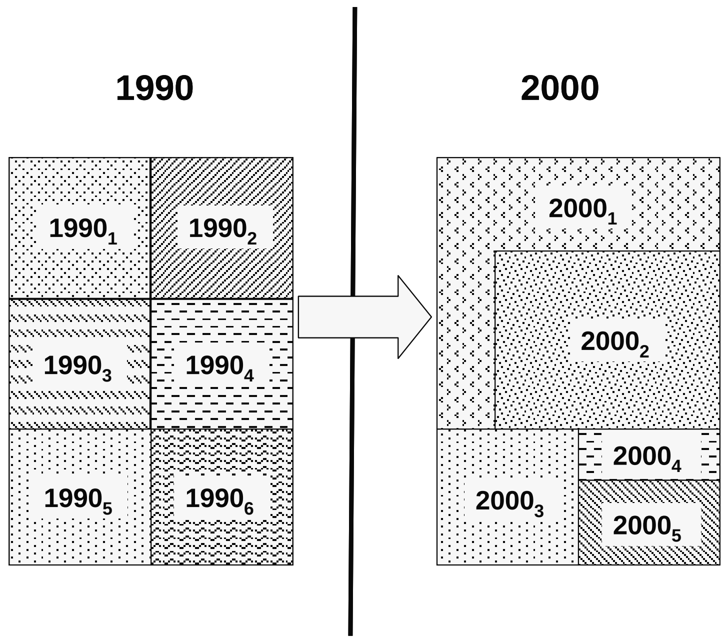
Panel F. Population Density (Low - Moderate Income Neighborhoods)

	Originate2004	Spread2004
BAccess04	49.858	-28.689
	(4.02) ***	(-2.51) **
R-square	0.78	0.41

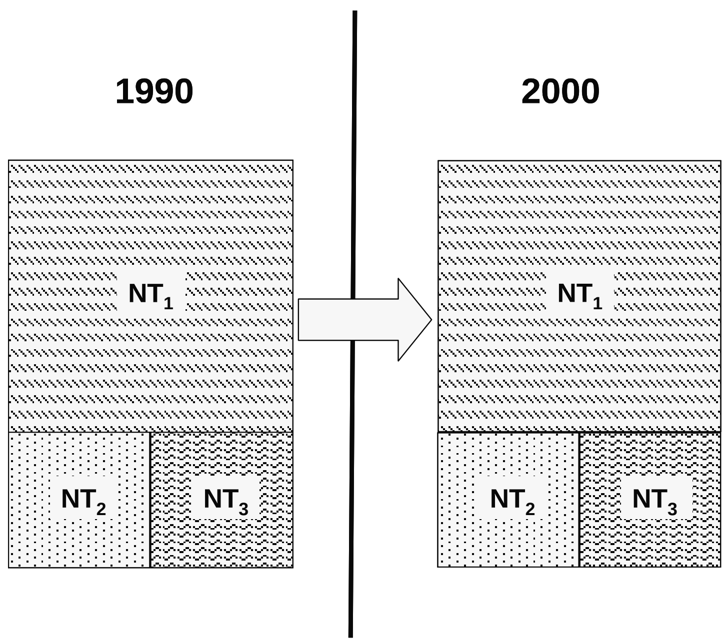
(***), (**), and (*) denote significant at 1%, 5%, and 10% level, respectively.

Figure 1. Creating Stable Census Tract Boundaries

A. Original Census Tracts



B. New Census Tracts



Appendix A. The computation of *MarketProfit99*

The idea behind the computation of the profitability of the banking market in and around a census tract is similar in spirit to the technique used by Amel and Liang (1997) in estimating the profitability of a market at the county level. I proceed as follows.

1. For each census tract, I find all bank branches within 20 miles of the census tract centroid.

2. I aggregate the deposits in those branches at the institution level if multiple branches belong to the same institution.

3. I choose the banks that raise 50% or more of their deposits from the area within 20 miles of a census tract. These are essentially small banks that do most of their business in the area. I will assume that their profitability is the profitability of doing business in that area.

4. To preclude the possibility of capturing a large bank that reports its aggregated deposits at its headquarters, I delete banks larger than $1 billion in total assets.

5. I calculate each remaining bank's profitability, Π_i, as

$$\Pi_i = \left(\frac{\textit{Interest Income from Loans}_i}{\textit{Total Loans}_i} - \frac{\textit{Interest Expense on Deposits}_i}{\textit{Total Deposits}_i} \right)$$

Note that this measure excludes the factors that may affect profitability at the institution level but are not related to the profitability of the market, such as the cost of other debt, which would be affected by the bank's access to capital markets, or non-interest expenses, which may be affected by the bank's operational efficiency.

6. *MarketProfit* is calculate as

$$MarketProfit = \frac{\sum_{i=1}^{n} D_i \Pi_i}{\sum_{i=1}^{n} D_i}$$

where n is the number of institutions and D_i is the total deposits of bank i.