Branches in Local Mortgage Markets

Lindsay E. Relihan*
The Wharton School
University of Pennsylvania

February 2017

Abstract

This paper studies the impact of branch presence on mortgage credit outcomes in the surrounding neighborhood using the density of nearby branch networks to instrument for actual branch presence. I find that lenders with branches lend more mortgages to borrowers in the surrounding neighborhood and that those operated by local lenders have the most positive impact for low socioeconomic-status borrowers. However, I show that branches disadvantage competing lenders by lowering the credit-quality of the competing lenders' applicant pool. This adverse selection causes an aggregate negative effect of branch presence on neighborhood mortgage outcomes.

*Contact by email at lrelihan@wharton.upenn.edu or by mail at 3000 Steinberg Hall-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104. I thank Todd Sinai, Fernando Ferreira, Gilles Duranton, and Joseph Gyourko for their continuous support of this project and Benjamin Keys and Anthony DeFusco for their helpful comments. I also appreciate the many useful discussions with other Real Estate faculty at The Wharton School and BEPP 900 seminar participants. Any errors are my own.

I Introduction I

The central goal of the Community Reinvestment Act (CRA) of 1977 was to increase access to credit for low socioeconomic-status borrowers. To that aim, the act encouraged lenders to open branches in more neighborhoods under the assumption that soft information about borrowers collected during in-person interactions between borrowers and loan officers would allow lenders to identify credit-worthy low socioeconomic-status borrowers who would otherwise be denied a mortgage (Lang and Nakamura 1993; Avery 1999; Essene and Apgar 2009). Building on this, the Interstate Banking and Branching Act of 1994 further incentivized banks to locate branches in such neighborhoods (Ludwig et al. 2009).[1] In the following years, thousands of new bank branches were established and the number of census tracts without any branch fell by 16 percent. Recent research suggests that the concurrent increase in mortgage credit access over the same period is partly due to the role branches play in the collection of soft information. For example, Favara and Imbs (2015) propose that soft information contributed to the 12 percent increase in state-level mortgage growth that they measure in states that removed branching restrictions after the passage of the 1994 act.

However, access to soft information about borrowers in the small business market, who have similarly low-quality hard information, is often found to decrease their access to credit (Broekner 1990; Petersen and Rajan 1995; Boot and Thakor 2000; Dell'Ariccia and Marquez 2004; Dell'Ariccia and Marquez 2006; Degryse and Ongena 2007; Presbitero and Zazzaro 2011; Gormley 2014). They attribute the decrease to an adverse selection problem between lenders with differential access to soft information. Intuitively, lenders with access to soft information are able to cream-skim the best borrowers with low-quality hard information. This leaves other lenders with a lower quality applicant pool, forcing them to either raise their lending standards or compete by specializing in a particular borrower type. Therefore, the total effect of soft information can be negative if the latter response dominates.

Given this widely supported result, it seems unlikely that branches would necessarily expand mortgage access to low-socioeconomic status borrowers through soft information. Outwardly, the

[1] This law included amendments to the CRA. It required regulators to evaluate banks' applications for out-of-state branch acquisitions and de novo out-of-state branching based on their performance in CRA evaluations. These evaluations are concerned with increasing banks' lending to low-income and minority groups in neighborhoods where they operate. This legislation, in effect, tried to mandate that new branches *would* increase access to mortgage credit for low-quality hard information borrowers.

ownership of a branch by a lender in a neighborhood would create a similar environment of asymmetric information with other lenders without branches. If that dynamic does exist, then federal policy intent on increasing soft information in the mortgage market could actually decrease credit for low-socioeconomic status borrowers.

The aim of this paper is to use a novel identification strategy and micro data to measure whether soft information, made available through branches, improves mortgage access for low-socioeconomic status borrowers in the surrounding neighborhood. The main challenge to finding a causal relationship is the endogeneity of branch location choice at the neighborhood level; for example, the choice of lenders to locate branches in neighborhoods with increasing mortgage demand would create a positive correlation between branches and mortgage credit. This correlation, then, may obscure the true relationship and lead to the erroneous conclusion that branches improve mortgage access.

My identification strategy is to build instruments for branch location choice based on a neighborhood's distance from a lender's pre-existing branch network.[2] I build a model of neighborhood mortgage markets to show that the validity of the instrument derives from the lenders' branch network optimization problem, in which economies of density in advertising and management lead lenders to establish new branches that are close to their pre-existing network (Berger 1997; Bos and Kool 2006; Felici 2008). Furthermore, the instrument is exogenous to credit access if the location of their pre-existing branches is uncorrelated with the contemporaneous borrower socioeconomic characteristics that also determine equilibrium mortgage credit.

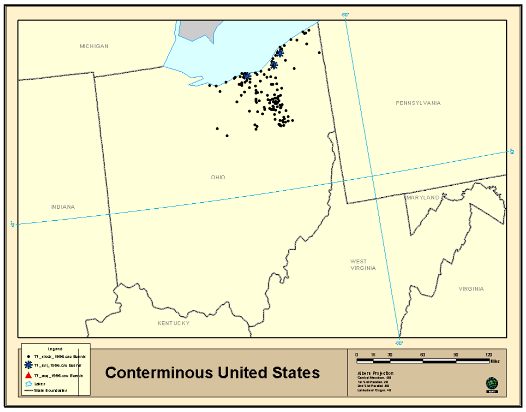
Figure 1 shows the typical pattern for growth of a lender's branch network, in which the network starts out as a small cluster of branches that gradually expands from its core.[3][4] Panel (a) shows the lender, FirstMerit Bank, in 1996 as a small local bank concentrated in the Cleveland area. That year, it opens three new branches along Lake Erie. Two years later in 1998 (panel (b)), it merges with another local bank and the combined entity continues to operate as FirstMerit. Then, in each successive panel it establishes new branches along the periphery and within its existing market. By

[2]This identification strategy is inspired by estimates in Holmes (2011) that found economies of density to be large and crucial for Wal-mart store locations. Goolsbee and Syverson (2004) also use network density as instruments in their study of airline pricing.

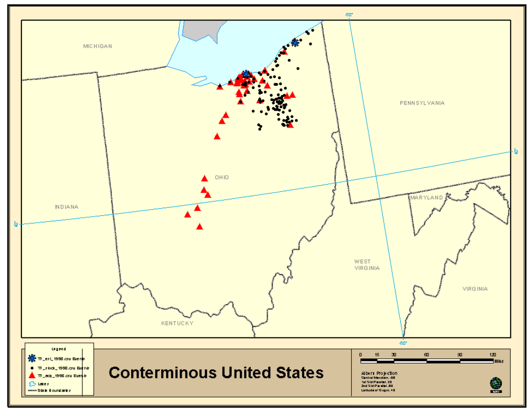
[3]This bank's growth pattern is not unusual. I chose it because the size of its network and its existence throughout my sample period make the pattern more easily discernible to the eye. Every other year from 1996 to 2008 is shown for compactness.

[4]For each panel, pre-existing branches are shown as small black dots, branches established in that year are shown as large blue stars, and branches acquired that year are shown as large red triangles.

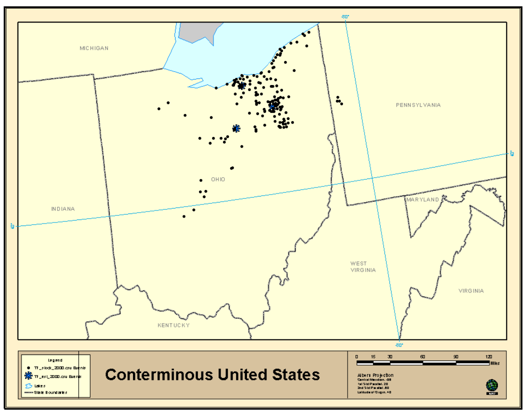
Figure 1: FirstMerit Bank Branch Network, 1996 - 2008.



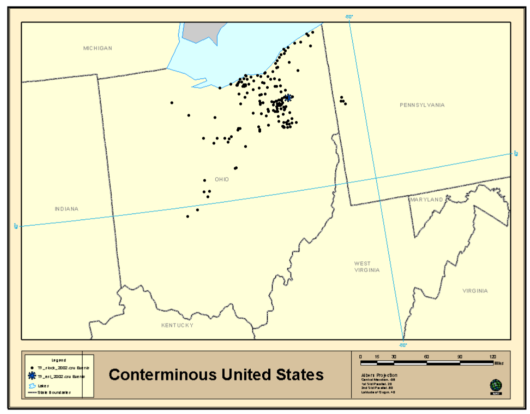
(a) 1996



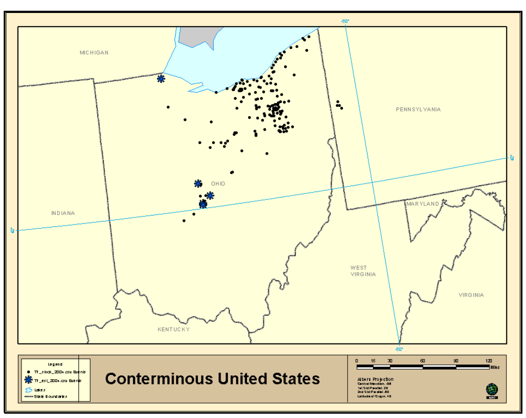
(b) 1998



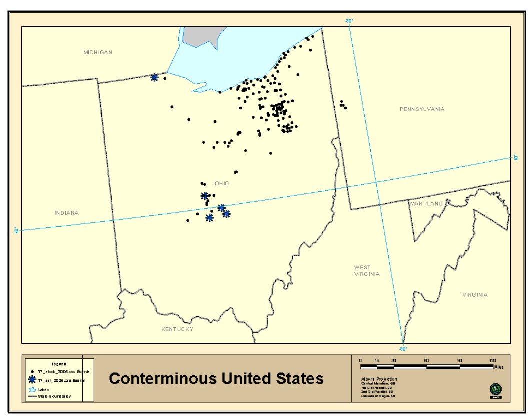
(c) 2000



(d) 2002



(e) 2004



(f) 2006

1 In each year, preexisting branches are shown as small black dots, newly established branches as large blue stars, and acquired branches as large red traingles. In a typical year, the bank adds a few new branches close to its existing branches.

2008, shown in panel (f), FirstMerit is a substantial regional presence with some branches stretching across state lines. This growth pattern is the bedrock of my identification strategy – networks that grow densely and slowly cannot easily respond to changing local economic conditions. It implies a low correlation between predicted branch location choice and neighborhood socioeconomics and allows identification of the causal effect of branches.

A second challenge is that soft information, by definition, is not directly measurable. To detect its use, I rely on two proxy variables that should be affected when lenders use soft information to make better loans: the percent sold to the secondary market and foreclosure rates three years after origination. These measure how profitable the lender itself thinks a loan will be and the realized credit quality of borrowers. In addition, I distinguish branches owned by small, local banks from those owned by large regional and national banks and those branches that specialize in mortgages from branches offering a general set of services. For these branches, the effects of soft information should be strongest because they have the most ability to collect and use it in their lending decisions.

Using this new identification strategy and detailed micro data, I assess the validity of two predictions from my model about lender-specific and aggregate lending behavior in mortgage markets with adverse selection. As in the first prediction of the model, individual lenders react negatively to a competitor's informational advantage due to the presence of their branch – by either raising lending standards or selling more mortgages to the secondary market. For instance, if a competitor to a small, local bank opens a branch, the average income of the borrowers who receive loans from the local bank increases by 0.9 percentage points. This implies tighter lending standards at the local bank and fewer loans for low-socioeconomics status borrowers. Therefore, even if the bank with the branch increases its lending to those borrowers, the aggregate effect of the soft information could be negative due to adverse selection. This is the second prediction of the model, and the results show that the aggregate effect of any branch type is, in fact, negative.

However, the model also indicates that the effect of adverse selection could be reduced if lenders have competitive advantages for lending to different types of borrowers. A lender's competitive advantage could be in either its cost of lending or the quality of its soft information signal from a branch. Then under such conditions, lenders could segment the market, rather than compete directly, and lend to more borrowers overall in equilibrium. Empirically, this matches the situation in which non-local banks have the advantage of lower capital costs, but local banks are better at collecting and using soft information. And indeed, the results show that the presence of a branch owned by a local bank has the least negative impact on low-socioeconomic status borrowers, most likely because they are able to cater to that borrower type.

My results suggest that national improvements in mortgage credit access subsequent to branching deregulation were not driven by increased access to soft information about borrowers surrounding new branches. In fact, that soft information lead to a decrease in credit access for low-socioeconomic borrowers, which ran counter to that specific mortgage market policy goal. However, my results do not necessarily negate the positive effect of branching deregulation on mortgage credit access found at more aggregate levels. For instance, branches could increase competition and lower mortgage rates for other borrowers (Calem 1998), increasing aggregate credit access, while still lowering access for specific low-socioeconomic borrowers. Furthermore, that outcome could be more efficient for the neighborhood mortgage market if fewer low credit quality borrowers receive loans.

Previous research on the link between soft information and neighborhood mortgage access has been hampered by the difficulty of controlling for endogenous branch location at the neighborhood level. Ergungor (2010) and Nguyen (2016) are the only other studies focused on measurement of the causal effect and use lagged demographics and branch closings due to bank mergers, respectively, to instrument for the presence of a branch in a neighborhood. They find a small, positive effect of the presence of branches on aggregate mortgage credit access. This study is different because it finds a negative effect using a new instrument based on the fundamentals of the lender optimization problem and examines the lender-specific responses that lead to the aggregate effect.

This paper relates to other research on the impact of branches and branching regulation (Jayaratne and Strahan 2006; Huang 2008; Beck et al. 2010; Kerr and Nanda 2010; Acharya et al. 2011), soft information in the mortgage market (Keys et al. 2010; Agarwal et al. 2011; Jiang 2013), the impact of CRA lending agreements (Schwartz 1998; Bostic and Robinson 2003), and the role of geography in economic outcomes (Moretti 2004; Giroud 2013; Carlino and Kerr 2015; Handbury et al. 2016). It shows that adverse selection also affects mortgage markets with asymmetric information, that the localized nature of soft information can create that asymmetry, and that instruments based on network density can be used to make causal estimates of the effects of branches.

II Data

The data used in my study include a much more detailed and nationally representative sample of loans and branches than any previous study on mortgage credit access. Most importantly, my data is the first to include information on mortgage brokers and non-bank lender branches, which now dominate a large portion of the origination market. Approximately 7,000 mortgage brokerage firms were operating in 1987 and originated around 20 percent of all mortgages. By 2003, the number of brokerage firms had risen to over 50,000 and they originated over 60 percent of all mortgages.[5] The inclusion of this data allow me to make a more accurate and comprehensive analysis of credit access within the prevalent industry structure.

A neighborhood in my study is defined as a census tract.[6] This is a particularly useful unit of observation, since census tract definitions try to keep population size and demographics somewhat constant while accounting for man-made and natural formations. Census tracts typically number less than 8,000 people with an target size around 4,000.[7] For each census tract in each year, I know the exact latitude and longitude location of branches within the tract and characterisitcs of almost every mortgage application and origination.[8] Observations in the merged data are either for a lender in a census tract in a year or a census tract in a year, depending on the context.

Mortgage loan origination data span 1994 - 2009 and come from the public use version of the Home Mortgage Disclosure Act (HMDA) database. HMDA was passed by Congress in 1975 and requires every lender satisfying any of a broad list of criteria to report every loan and a set of its

[5] "Mortgage brokers fall on tough times." USA Today, Web. 31 Sept 2007.

[6] Due to census tract definition changes between the 1990 and 2000 censuses, I form consistent geographic units with the census tract relationship file. Census tracts from 1990 and 2000 are assigned to the smallest geographic area that can be held constant across the censuses. This results in an average "tract" size that is slightly larger than census tracts in either the 1990 or 2000 files.

[7] Median census tract size is about 2 square miles. I drop census tracts in the top 5 percent of the size distribution, as these are not suited to my spatial analysis. For more information, see `https://www.census.gov/geo/reference/pdfs/GARM/Ch10GARM.pdf`

[8] The are several minor datasets I do not detail. Shape files for census tracts for the 1990 and 2000 Census definitions come from the National Historical Geographic Information System. Relationship files for standardizing tract definitions across Census years come from the United States Census Bureau. Reports of Condition and Income (Call Reports) from the Federal Financial Institutions Examination Council were used in matching regulatory identification numbers between the FDIC and HMDA data. The 2009 5-year American Community Survey dataset was used to collect one time demographic information on census tracts. Property deeds records collected by DataQuick were used to calculate the 3-year cumulative foreclosure rate. Their geographic coverage is less extensive than HMDA and concentrated in more populated areas. Data Quick does not have a common identifier with the mortgage loan or branch data, and so I can only calculate foreclosure rates at the neighborhood and not the neighborhood by lender level.

characteristics to a central repository.[9] The sample of loans contains both purchase and refinance loans, first liens and junior liens. The estimated coverage of HMDA data for first-lien purchase originations is around or over 90 percent for my sample period (Scheessele 1998). My own estimates for other loan types show similar coverage.[10]

Data on mortgage broker and non-bank lender branches come from the Nationwide Mortgage Licensing System (NMLS) maintained by the Conference of State Bank Supervisors.[11] The NMLS data contain licensing information at the company, branch, and loan officer level. A snapshot from the earliest, nationally comprehensive data provides the stock of these licenses since 2008 at each level, nationwide. There are several limitations to the NMLS data. Due to varying licensing regulations at the state level before 2008, the existence of each branch and its dates of operation are calculated through a combination of office licenses and individual licenses for individuals working at each location.[12][13] A further limitation is that branches that closed before 2008 are not in this database. In comparison to outside estimates of the size of this industry, the 2008 NMLS data contain about half the number of these companies in operation at the height of the housing boom.[14]

A number of stand-alone lenders have large retail operations.[15] The distribution of NMLS network size is presented graphically in Figure 2. From 1994 to 2009, the size of the largest network grows from around 200 branches to over 900 branches, shown by the lightest dashed line. There are also many small lender branch networks. The average size, shown by the green dashed line, goes from less than 1.5 branches to almost 2 branches by 2009.

The rise in importance of mortgage brokers and non-bank lenders to the mortgage market is

[9] For more information on criteria see the Federal Financial Institutions Examination Council's report *A Guide to HMDA Reporting: Getting it Right!* at `http://www.ffiec.gov/hmda/pdf/guide.pdf`

[10] HMDA refinance loans and junior liens were matched to public record files compiled by DataQuick

[11] This database was created in the wake of the recent housing crisis by the Secure and Fair Enforcement for Mortgage Licensing Act (SAFE) of 2008. The SAFE Act created national licensing standards for entities and individuals involved in the mortgage origination process and requires state agencies to provide such information to the NMLS database.

[12] See the excellent compilation of state regulations from 1996 - 2006 from the Federal Reserve Bank of Minneapolis at `http://www.minneapolisfed.org/publications_papers/pub_display.cfm?id=4983&`

[13] Branch licenses only list the issue date of the license, which may only be issued once states require branch licensing. Individual licenses list both the issue date and the date at which the employee began working at her current location. For some states with strict licensing requirements well before 2008, such as California, I am confident of calculating the correct establishment date. For other states, such as Alaska, which had no licensing standard of any kind before 2008, the establishment date is less precise. Most states fall somewhere in the middle of these two extremes.

[14] Adding data for companies before 2008, assuming it did and continues to exist, would require separate Freedom of Information Act requests to each individual state regulatory agency. Therefore, despite its limitations the NMLS data is the best source for information on the location and operation of non-bank branches.

[15] For example, American Pacific Mortgage operates almost 300 branches in my data.

highlighted by Figures 3 and 4. Figure 3 shows that in 1994, my data contain less than 4,000 unique firms. That number climbs to over 20,000 by 2009. This steady increase through the housing bust reflects the survivorship bias in my sample of these firms.[16] Figure 4 shows how many branch locations mortgage brokers and non-bank lenders owned at each point in my sample. From 1994 to 2009 these firms went from operating 5,000 branches to almost 40,000 branches.

Bank branch networks and the number of bank branches exhibit similar trends as in the NMLS data. Data on bank branches from the Federal Deposit Insurance Corporation's (FDIC) Summary of Deposits database from 1994 - 2009 are summarized in Figures 5 and 6.[17] The database includes information on every bank branch with federally insured deposits. The information at the branch level includes address, branch type, specialty, date established, and date acquired. Information on the bank, bank holding company, and regulator of the branch is also included. Through available identifiers, I am able to match each bank directly to the loans it originates, but unfortunately not to the branch of origination.

In my analysis, I classify branches both by type of lender and the size of the network to which they belong. There are four types: general service bank branches, bank branches specializing in mortgages, broker branches, and non-bank lender branches. Bank branches are classified as mortgage specialists if they are classified that way in the FDIC data at least half of the time they are in operation. About 26,000 branches, or 8 percent of bank branches, receive this classification. A branch is defined as belonging to a mortgage broker if that company is ever issued a license from a state that licenses mortgage brokers as distinct from mortgage lenders (about half of states).[18] About two-thirds of the NMLS branches are then classified as mortgage broker branches. For bank branches, I also make a distinction by network size. Local branches are branches with fewer than 37 branches when I first observe them – this is the 99th percentile for network size in 1994. About half of bank branches are local.[19]

Table 1 provides yearly summary statistics on the mortgage characteristics in census tracts.

[16]Firms with the poorest mortgage practices and located in the hardest hit areas are likely underrepresented due to failure before 2008.

[17]Under the Interstate Banking and Branching Efficiency Act (IBBEA) of 1994, banks for the first time were allowed to widely own and acquire other banks across state lines and to operate and open new branches across state lines. For details see Johnson and Rice (2007).

[18]This likely causes some miss-classification of NMLS lenders. Results are not sensitive to other classification rules.

[19]Mortgage brokers and non-bank lenders are not separated by network size – 90 percent would qualify as local.

The first two columns show that during the boom, lenders were more likely to sell originations into the secondary mortgage market. Changes in the 3-year foreclosure rate are shown in columns 3 and 4. The median here reflects the cycle of boom and bust, but the standard deviation highlights the extreme variance in foreclosure rates across neighborhoods. Columns 5 - 8 show that the reported income of borrowers during the boom fell, although more of them received conventional loans (as opposed to government financed mortgages like those from the VA or FHA).

Table 2 provides summary statistics for the stock of branches and mortgage demand and supply by year for the 55,000 census tracts in my sample. The first two columns examine the stock of branches within a census tract. From this, it is easy to see that the vast majority of census tracts have a small number of branches. Over time, the average and standard deviation of the stock of branches goes up, indicating that new branch openings are concentrated in a subset of the census tracts. The remaining columns show statistics for log originations, log applications, and the percent of applications denied. They show a pattern consistent with recent housing history, in which credit access was reduced slightly during the downturn around 2000 and severely during the recent crisis. Overall, these two tables match many of the well documented statistics about the mortgage market over the recent cycle (Mian and Sufi 2011; Ferreira and Gyourko 2011).

III Model

In this section, I develop a two-lender model of neighborhood mortgage credit demand and supply to show how asymmetric soft information can affect lenders' credit standards and motivate my choice of empirical specifications.[20] The key insight from the model is that imperfect information about borrower quality drives a wedge between the cost of the loan to the lender and the mortgage rates available to borrowers. I call this wedge the information markup. Adverse selection due to asymmetric information changes the information markup charged by each lender, leading to changes in lender specific and aggregate credit access. I combine this model of mortgage credit

[20] My model adapts prior theoretical research (beginning with Stiglitz and Weiss (1981)) to a mortgage market with imperfect information. The search and negotiation behavior of consumers when search is costly has been well explored (Burdett and Judd 1983; Wolinsky 1987; Bester 1988). A strategy of sequential search with a reservation price and auction is close to the optimal strategy described in McAfee and McMillion (1987). In the mortgage market, Allen et al. (2014) use a similar model to explain dispersion in mortgage interest rates. Finally, I follow tradition in the small business literature and use signals of borrower credit quality to proxy for information quality (Broeckner 1990; Hauswald and Marquez 2006).

demand and supply with a model of branch network optimization to show that lenders maximize profit by reducing network costs through higher density, the basis for my identification strategy.

In the first stage of the model, each lender decides whether to operate a branch in a neighborhood versus lend without one. Subsequently, households sequentially search for a mortgage from lenders originating loans in their neighborhood. Once contacted, lenders and borrowers receive a signal of a borrowers' credit quality. Lenders with a branch produce a higher quality signal due to the availability of soft information. Lenders then decide whether to make an initial mortgage offer to the borrower. If an offer is made, the borrower can accept or pay the search cost to seek other offers. Households with multiple offers hold an auction. Household search concludes once every household has a mortgage or has left the market.

Households: Consider a neighborhood with a unit mass of households, indexed by i, whose utility depends on housing consumption, $M_i \epsilon \{0,1\}$, and non-housing consumption, $c_i > 0$. Households are risk neutral and identical in their income, I, marginal utility of housing, $\mu > 1$, and mortgage search costs, ρ. They only differ in their creditworthiness, $\theta_i \epsilon \{\theta_h, \theta_l\}$, in that high credit quality households of type θ_h always repay their mortgage and low credit quality households of type θ_l always default.

Households in a neighborhood choose whether to buy one of the identically and perfectly elastically supplied houses available at a normalized price of 1. Households must take out a mortgage from a lender operating in that neighborhood to completely finance the purchase of the house. Households who do not take out a mortgage only purchase non-housing consumption.

The household utility maximization problem takes the form:

$$\max_{0 \le c_i \le I_i, M_i = 0,1} c_i + \mu M_i \tag{1}$$

subject to the budget constraint:

$$I - c_i - r^e_{\theta_i} M_i \ge 0. \tag{2}$$

where $r^e_{\theta_i}$ is the effective interest rate they face after n searches for a mortgage. Search for a mortgage is conducted sequentially, with households matched randomly to their initial lender and to the remaining lenders during any additional search.[21] Consumers will accept the mortgage contract if condition (2) holds, in equilibrium they do not expect to receive a better rate by searching further, and paying for the mortgage is utility maximizing:

$$(\mu - r^e_{\theta_i})\mathbb{1}\{M_i = 1\} > 0. \tag{3}$$

Lenders: On the supply side, lenders (indexed by j) observe the I, μ, and ρ of the neighborhood. They also observe the quantity of θ_h and θ_l households in the neighborhood, but not the θ_i of an individual household. Lenders face a cost of supplying a loan, $c^j_{\theta_i} > 0$, and offer interest rates such that the expected profits are non-negative for each borrower. Lenders have access to a screening technology that produces a signal, η^j, that a borrower is of type θ_h or θ_l. The signal, however, is imperfect. Let $q^j = Pr(\eta^j = \theta_h|\theta_i = \theta_h) = Pr(\eta^j = \theta_l|\theta_i = \theta_l)$ be the quality of the signal. Lenders have the option of opening a branch in the neighborhood, which improves their screening technology through the availability of soft information. Thus, q^j is a function of branch presence such that $q^j(b = 1) > q^j(b = 0)$. The opening of a branch incurs a cost, $\tau^j v^j$, where v^j is the cost of operating the branch and τ^j is the distance to the existing branch network. Conditional on their branch entry decision, lenders will offer interest rates that maximize expected revenue, meaning:

$$\max_{0 < r^j_{\theta_i} \leq 1} Pr(\theta_i = \theta_h|\eta^j_i = \theta_i) r^j_{\theta_i} \tag{4}$$

such that the borrowers accept (meaning (2) and (3) hold) and expected revenue is greater than the cost of funding the loan:

$$Pr(\theta_i = \theta_h|\eta^j_i = \theta_i) r^j_{\theta_i} \geq c^j_{\theta_i} \tag{5}$$

Mortgage Negotiation: Prospective borrowers and lenders participate in a game with the following stages. The game is solved by backward induction.

[21] This choice is due to the substantial search frictions in the mortgage market – consumers are generally confused about the origination process and comparing options requires contacting each lender individually (Woodward and Hall 2010 and 2012; Allen et al. 2014). Furthermore, a recent survey from the Consumer Financial Protection Bureau found that half of borrowers apply to only one lender.

(1) Lenders decide if they should open a branch in the neighborhood, paying cost $\tau^j v^j$ to gain access to a higher quality signal of household type.

(2) Households are matched randomly with one of the lenders and they each receive the same signal of household type. Each Lender makes an initial offer to each borrower such that expected profits are non-negative.

(3) Households decide based on their type signal and initial offer if they should search for an additional offer from another lender. Those that do search pay the search cost.

(4) Households with multiple offers receive a new offer of the lowest possible interest rate from the lowest cost lender. If both lenders have the same lowest offer, households choose one randomly.

(5) Continue this way until all borrowers receive a mortgage or have left the market.

Solving the Simple Game: I consider the simplest version of the model, in which there are only two lenders operating in a neighborhood with equal costs of supplying a loan to both borrower types and opening a branch. The full details of solving this simple model can be found in Appendix C.

In the two cases of either no lender operating a branch or two lenders operating branches, there is no information asymmetry. Households who receive a signal that they are type θ_h know that they will not receive a better interest rate through searching, while those with a signal of type θ_l may search if the expected gains are high enough, i.e. when search costs are low, the marginal utility of owning is high, or incomes are high. Using Bayes' theorem I can solve for the lenders' lowest rate they are initially willing to offer borrowers with $\eta^j = \theta_i$:

$$\underline{r}^j_{\theta_h} = \begin{cases} c[1 + \frac{\theta_l}{\theta_h}\frac{1-q}{q}\frac{1+q}{2-q}] & \text{if } \eta_i^{-j} = \theta_l, n = 1 \\ c[1 + \frac{\theta_l}{\theta_h}\frac{1-q}{q}] & \text{otherwise} \end{cases} \tag{6}$$

$$\underline{r}^j_{\theta_l} = \begin{cases} c[1 + \frac{\theta_l}{\theta_h}\frac{q}{1-q}\frac{1+q}{2-q}] & \text{if } \eta_i^{-j} = \theta_l, n = 1 \\ c[1 + \frac{\theta_l}{\theta_h}\frac{q}{1-q}] & \text{otherwise} \end{cases} \tag{7}$$

The second term inside the brackets in equations (6) and (7), is the information markup and captures the negative relationship between information quality and rates. With full information about borrowers, lenders could offer rates at cost for θ_h households and deny loans for θ_l households. When household type is uncertain, the information markup increases with the proportion of type θ_l borrowers in the neighborhood and further increases if they are expected to apply to both lenders (when $\eta_i^{-j} = \theta_l, n = 1$). But as screening technology improves, households with $\eta^j = \theta_h$ receive lower rates and households with $\eta_i^j = \theta_l$ receive higher rates. Of course, the lender may be able to charge borrowers interest rates above the minimum and increase profits if I, μ, or ρ are high.

In the case in which only one lender, say L1, decides to operate a branch, that lender will have better information about borrower type than the other lender, say L2. Here, borrowers of type θ_l who are matched with the L1 may decide to search and receive a lower offer from the less informed lender, L2, who would be less certain that they are of type θ_l. And borrowers of type θ_h who are initially matched with L2, may also decide to search and receive a lower offer from L1, who would be more certain that they are of type θ_h. In this case, each lender's lowest rate they are initially willing to offer will be:

$$\underline{\mathrm{r}}^1_{\theta_h} = \begin{cases} c[1 + \frac{\theta_l}{\theta_h}\frac{1-q^1}{q^1}\frac{1-q^2}{1+q^2}] & \text{if } \eta_i^2 = \theta_h, n = 1 \\ c[1 + \frac{\theta_l}{\theta_h}\frac{1-q^1}{q^1}] & \text{otherwise} \end{cases} \tag{8}$$

$$\underline{\mathrm{r}}^1_{\theta_l} = \begin{cases} c[1 + \frac{\theta_l}{\theta_h}\frac{q^1}{1-q^1}\frac{2-q^2}{1+q^2}] & \text{if } \eta_i^2 = \theta_h, n = 1 \\ c[1 + \frac{\theta_l}{\theta_h}\frac{q^1}{1-q^1}] & \text{otherwise} \end{cases} \tag{9}$$

$$\underline{\mathrm{r}}^2_{\theta_h} = \begin{cases} c[1 + \frac{\theta_l}{\theta_h}\frac{1-q^2}{q^2}\frac{1+q^1}{2-q^1}] & \text{if } \eta_i^1 = \theta_l, n = 1 \\ c[1 + \frac{\theta_l}{\theta_h}\frac{1-q^2}{q^2}] & \text{otherwise} \end{cases} \tag{10}$$

$$\underline{r}^2_{\theta_l} = \begin{cases} c[1 + \frac{\theta_l}{\theta_h}\frac{q^2}{1-q^2}\frac{1+q^1}{2-q^1}] & \text{if } \eta_i^1 = \theta_l, n = 1 \\ c[1 + \frac{\theta_l}{\theta_h}\frac{q^2}{1-q^2}] & \text{otherwise} \end{cases} \tag{11}$$

Searching amounts to multiplying the information markup by an additional term. In the case of L1, when households with $\eta_i^2 = \theta_h$ are part of its applicant pool, it raises the average quality and thus lowers the rates it needs to charge households with both signal types. Conversely, when households with $\eta_i^1 = \theta_l$ are part of L2's applicant pool, it lowers the average quality and necessitates a rise in interest rates for households with both signal types. This is a clear illustration of how asymmetric soft information can lead to an adverse selection problem.

Given the number of branches, the overall amount of credit in neighborhood i, is determined by the rates offered by each lender as a function of θ_h, θ_l, I, μ, and ρ (See Appendix C for the equation which gives the equilibrium amount of credit). If $\underline{r}^j_{\theta_h}$ and $\underline{r}^j_{\theta_l}$ are low enough, then all borrowers will want a mortgage. If both are too high, then there will be no mortgage demand. And if only $\underline{r}^j_{\theta_h}$ is low enough, then only borrowers who receive at least one signal that they are of type θ_h will demand mortgages.

The most important fact highlighted by the model is that the soft information acquired by lenders through branches does not unambiguously increase credit access for all types of borrowers. Moving from no branches to one or two branches, unambiguously lowers rates for θ_h borrowers, but raises rates for θ_l borrowers. These offsetting effects could cause the presence of a branch to lower the aggregate number of borrowers who receive credit.

Branch Entry: It remains to determine which lenders will open a branch in the neighborhood. Denote $\pi_e^j(b^j, b^{-j})$ as the expected profit from mortgages of lender j given its own and competitor $-j$'s entry decision. Lenders face the following game matrix to determine equilibrium entry:

		L1	
		$Branch$	Not
L2	$Branch$	$\pi_e^1(1,1) - \tau^1 v, \pi_e^2(1,1) - \tau^2 v$	$\pi_e^1(0,1), \pi_e^2(1,0) - \tau^2 v$
	Not	$\pi_e^1(1,0) - \tau^1 v, \pi_e^2(0,1)$	$\pi_e^1(0,0), \pi_e^2(0,0)$

The payoffs reflect that building a branch in neighborhood i will have a higher payoff in areas with higher expected demand, more high quality borrowers, higher incomes, higher marginal utility from housing, and higher search costs. Such neighborhoods allow lenders to sell more loans and at rates above their expected cost. The payoff matrix also shows that profits are higher when neighborhood n is closer to lender j's network, due to the assumption that cost increases with distance τ^j.

Predictions: The model gives a set of predictions about where branches should locate and their effect on mortgage credit if that effect works through the use of soft information. The first two focus on detecting the presence and effect of adverse selection and the third on causal identification:

(1) Lenders prefer to locate branches in locations that fit into a dense network of branches, have less competition, and more profitable borrowers.

(2) Lenders operating branches will extend more credit to borrowers with low-quality hard information. Lenders without a branch will lend less to these borrowers when their competitors operate a branch.

(3) The aggregate credit response to a branch is ambiguous, and depends on the combined positive and negative responses of individual lenders.

IV Identification and Empirical Approach

The model highlights the central identification challenge of isolating the causal effect of branch presence on credit outcomes. While branches may affect credit access through the availability of soft information, high credit demand and potential profits themselves may encourage entry. Therefore, any simple regression of credit outcomes on branch presence will necessarily be biased, most likely toward increased credit access. This can be seen by writing a linearized version of the quantity of credit, Q_{nt}, in equilibrium for neighborhood n at time t:

$$Q_{nt} = \beta_0 \pi_{nt}(I, \mu, \rho, \theta_h) + \beta_1 Branch^1_{nt} + \beta_2 Branch^2_{nt} + u_{nt} \tag{12}$$

where $\pi_{nt}(I, \mu, \rho, \theta_h)$ is the profitability of neighborhood n, $Branch^j_{nt}$ is an indicator for whether or not lender j has a branch in neighborhood n at time t, and u_{nt} is an iid error term. The profitability of the neighborhood is observable to the lenders, but not to the econometrician. Being unable to control for it, a regression which omits it will necessarily be biased, as $e_n = u_n + \beta_0 \pi_{nt}(I, \mu, \rho, \theta_h)$ will covary with each $Branch^j_{nt}$.

My identification strategy relies on the property that lenders are more profitable when they build dense networks of branches to minimize the operating costs, $\tau^j_n v^j$. Therefore, neighborhood distance from the existing network will be an important factor in the branch entry decision.

$$Branch^j_{nt} = \Gamma_0 \pi_{nt}(I, \mu, \rho, \theta_h) + \Gamma_1 Branch^{-j}_{nt} + \Gamma_2 \tau^j_{nt} + v_{nt} \tag{13}$$

I argue that the distance from existing branch networks is a strong predictor of branch presence in a neighborhood, but is not related to the credit conditions of that neighborhood. More formally, that $Cov(v^j_{nt}, e_{nt}) \approx 0$.[22] This is because the vast majority of lenders in my sample were founded well before my sample period, based on the economic environment of that time and not today. Subsequently, it would be very difficult for lenders to quickly respond to changing economic

[22] It is unlikely that this covariance is exactly zero due to long run persistence in local economic growth, foresight by lenders, etc., but it is assumed to be small.

conditions due to the large financial and regulatory fixed costs of branch network adjustment.[23][24]

To illustrate this point, the data in my sample show that in a year a lender opens an average of only 0.3 branches, about 9 percent of their total stock. And, on average, only 15 percent of lenders open one or more branches in any year. Closings are even more rare. On average, 0.13 branches are closed per lender each year, about 0.7 percent of their existing stock. This slow adjustment process inhibits lenders' ability to open and close branches based solely on short-term mortgage profitability. Thus, distance between the neighborhood and lender j's existing network of branches can serve as a valid instrument for the presence of lender j's branch in neighborhood n.

Furthermore, I can show that distance predicts branch presence through my entire sample, not just branch reorganization after mergers. If I exclude years after 2002 (when a wave of bank mergers began) or major mergers during that time period, my instruments are actually stronger (see Appendix B). This suggests that my instruments capture a more fundamental feature of the branch location decision that instead supports instruments based on bank mergers, rather than the other way around. Specifically, that the openings and closings of branches post-merger are exogenous to local economic outcomes because the combined firm must re-optimize the entire branch network to fully exploit economies of density. However, my instrument allows me to study the effect of any branch and not just those affected by bank mergers.

Apart from the endogeneity of branch location, the causal interpretation of my results is threatened by other neighborhood features that affect neighborhood mortgage activity and branch location decisions. The panel structure of my data allows me to circumvent any that are fixed at the neighborhood, neighborhood by lender and/or year levels using fixed effects. This demeaning removes the omitted variables bias from any factor that remains fixed at the fixed effect level and affects mortgage outcomes.[25] My coefficients are then identified off of the deviations from the mean

[23] Bancography, a consulting firm specializing in branch planning, has conducted several surveys of the start-up costs for a new branch. Physical capital costs for a free-standing branch, including construction, equipment, and furniture, typically range from \$1 - \$1.5 million dollars. Other start-up costs include land and the hiring and training of new staff. For more information, see http://www.bancography.com/downloads/Bancology0803.pdf

[24] The requirements for opening, relocating, or closing a bank branch depend on the bank's regulatory agency. To open a new branch, a bank submits an application that often requires such information as compliance with the Community Reinvestment Act, an environmental impact statement, and satisfaction of local zoning regulations. To close a branch, a bank typically must give at least 90 days notice to the regulatory agency and its customers.

[25] Of course, fixed effects do not remove the bias from omitted variables that vary over time and are correlated with both lender branch presence and mortgage outcomes and are unrelated to branch network optimization. Ideally, I would control for the variables that I show correlate with branch location decisions, but these census tract level statistics are only available in decadal census years. To the extent that census tract population, income, and other

of a mortgage outcome for a neighborhood (or lender in a neighborhood) in a year from deviations in the mean predicted stock of branches (or branches owned by a lender) in a neighborhood in a year. Given the thousands of fixed effects that this sometimes entails, my estimates will be subject to attenuation bias and should be interpreted with this in mind.

Beyond these causal concerns, attributing my estimates to the presence of soft information is its own identification challenge. In the mortgage market, soft information is derived from features – the stability of the borrower's income, her character, the quality of the mortgage collateral, etc. – that are inherently difficult to quantify (Stein 2002; Keys et al. 2010; Agarwal et al 2011). Therefore, I rely on a pair of proxy variables that should be affected by the use of soft information by lenders to detect its influence. Specifically, I measure the percent of originations sold to the secondary market, because lenders may sell more of their loans (avoiding future losses from defaults) if they have low confidence in borrower credit quality (Keys et al. 2010). In addition, I measure cumulative foreclosure rates three years after origination as an ex post measure of the use of soft information, under the assumption that higher quality mortgages will default less often. Then, I can use changes in these proxy variables to infer the effects of soft-information on low socioeconomic status borrowers, as characterized by their income and qualification for conventional loans.[26]

In addition to the use of proxy variables, I work toward being able to attribute my findings to soft information by focusing on a set of census tracts with a very clear information structure that matches my model. These are census tracts that have zero branches in 1994 and either never experience a branch opening or transition from zero to one branch. This means that at the beginning of my sample, in these tracts, no lender has an informational advantage and, in those where one branch is later established, one lender clearly gains a soft informational advantage over all others. Then, any change in my outcome variables due to a change in the information structure between lenders is most likely caused by the soft-information collected by that branch.[27] Focusing on this sub-sample has additional advantages in that (1) they are predominately rural and less likely

features are relatively invariant in the short term, much of their effect should be captured in the fixed effects.

[26]To qualify for a conventional loan, a borrower must meet certain eligibility and financing requirements that may be difficult for low-socioeconomic status borrowers to meet. For more on these requirements see https://www.fanniemae.com/content/eligibility_information/eligibility-matrix.pdf

[27]Studying different combinations of lenders with and without branches is also practically difficult, because there are not enough of these transitions to detect any statistically significant results. By far the most common transition is from 0 to 1 branch.

to be affected by nearby dense branch networks, (2) best match the simple model on which my predictions are based, and (3) are the type of neighborhoods that policies encouraging branching are intended to help. Figure 1 gives the location of the census tracts.

Turning to my empirical approach, I capture the density of a lender's branches around a neighborhood by the number of a lender's branches within rings of distance from a neighborhood: from 5-10, 10-20, 20-50, and 50-100 miles away. Likewise for competitor density. Overall distance is captured by two features: neighborhood distance from the geographic center of the branch network and the standard distance between branches in the network. Standard distance is a measure of geographic spread; networks which are more spread out, conditional on branch location, can build new branches that are farther away without lowering network density.

At the aggregate tract level, I use instruments that try to capture both the density of nearby branches and the density of individual lenders. To this goal, I use the total number of branches within the same rings of distance from the census tract centroid, the number of lenders operating at least 5 branches within those rings, and the number of network centers that are located within each of those rings. For instruments for a particular branch type, I amend this by using the number of lenders operating at least 2 branches of that type in each ring. I use this smaller number since each branch type is less common than any individual branch.[28]

I now lay out the estimating equations that will build evidence for the effect of soft information on local mortgage markets. They are based on the predictions from the model of section IV and my strategies for dealing with the identification challenges posed by lender branch endogeneity and attribution to soft information.

Hypothesis 1: Endogeneity of branch location choice.

The first hypothesis is that lenders prefer to locate in neighborhoods with more profitable borrowers and lower network costs from branch network density. This provides evidence for the existence of the endogeneity problem and for the relevance of my instruments. I run equations of the form:

$$Branches_{itj} = \Sigma_{n=1}^{N} \beta_n NetDist_{nitj} + \Sigma_{n=1}^{M} \beta_m CompDist_{mit-j} + \lambda X_i + \phi_j + \gamma_t \tag{14}$$

[28] In Appendix B, I provide a version of my results where I construct a predicted number of branches for each lender in a census tract and then use the sum as the predicted aggregate number of branches in a census tract. The results are widely consistent with my main set of tract level instruments.

where $Branches_{itj}$ is the number of branches operated in neighborhood i at time t by lender j, the coefficient β_n measures the coefficient on $NetDist_{nitj}$ (one of N measures of distance to the lender's own branch network), the coefficient β_m measures the coefficient on $CompDist_{mitj}$ (one of M measures of distance to the competing branch networks), and λ is a vector of coefficients on the set of X_i demographics that correspond to income, search costs, and the propensity to demand a mortgage. The X_i variables do not vary over time, and so this specification does not include fixed effects for each census tract, but does include lender and year fixed effects, ϕ_j and γ_t. Due to this data limitation, the primary purpose of this specification is to highlight the endogeneity of branch location choice.

First-stage estimates.

The instrument lender-specific versions that do serve as the first stage are given by:

$$\widehat{Branches}_{itj} = \Sigma_{n=1}^{N} \beta_n NetDist_{nitj} + \Sigma_{m=1}^{M} \beta_m CompDist_{mit-j} + \alpha_{ij} + \gamma_t \tag{15}$$

$$\widehat{Branches}_{it-j} = \Sigma_{n=1}^{N} \beta_n NetDist_{nitj} + \Sigma_{m=1}^{M} \beta_m CompDist_{mit-j} + \alpha_{ij} + \gamma_t \tag{16}$$

where the α_{ij} are lender by census tract fixed effects.

The first stage for the aggregate number of branches in a tract, regardless of lender, is the predicted number of aggregate branches given by:

$$\widehat{Branch}_{it} = \Sigma_{p=1}^{P} \beta_p NetDist_{pit} + \alpha_i + \gamma_t \tag{17}$$

where β_p is the coefficient on $NetDist_{pit}$ (one of P measures of distance to any lender branch networks).

For regressions that look at a specific branch type or multiple lenders operating branches, the equations are analogous to (15) - (17) with $\widehat{Type}_{it}$ or $\widehat{Competitors}_{it}$ substituted for $\widehat{Branch}_{it}$ and independent variables adjusted accordingly.

Hypothesis 2: Lender-specific effects.

The second hypothesis is that lenders operating a branch in a neighborhood are able to screen out and lend to more high-quality borrowers with low-quality hard information because they now have access to their soft information. Additionally, when a competitor bank opens a branch, banks lend less to low-quality hard information borrowers and retreat to borrowers with higher hard information quality. For the subsample of banks with branches (the only lenders for whom I can match mortgage information) I run equations of the form:

$$Y_{itj} = \beta_1 \widehat{Branches}_{itj} + \beta_2 \widehat{Branches}_{it-j} + \alpha_{ij} + \gamma_t \tag{18}$$

where Y_{itj} is a mortgage outcome, $Branches_{it-j}$ is the number of branches in neighborhood i at time t operated by other other $-j$ lenders.

Hypothesis 3: Aggregate effects.

The third and final hypothesis is that presence of a branch has an ambiguous effect on overall credit access, and may depend on the type and size of the lender operating the branch (results from the modification of the model to include these variations are found in Appendix C). For the subsample containing tracts before or while they have their first branch, I estimate equations of the form:

$$Y_{it} = \beta_1 \widehat{Branch}_{it} + \alpha_i + \gamma_t \tag{19}$$

where Y_{it} is the aggregate mortgage outcome from neighborhood i at time t and α_i and γ_t measure census tract and year fixed effect.

V Results

Hypothesis 1:Endogeneity of branch location choice.

Evidence of the branch location endogeneity problem is useful for showing that an identification strategy is necessary for finding causal effects and for setting expectations about the direction of bias in OLS estimates. The evidence presented here is that lenders choose branch locations based on the profitability of local borrowers, even after controlling for the location of local branch networks. The results in Table 3 show that lenders have a strong preference for locating branches in census tracts with more households, higher household income, and prefer whiter neighborhoods – which also tend to be wealthier.[29] There is also a strong negative correlation with owner-occupancy, suggesting that lenders prefer to locate branches in areas with more potential mortgage demand. Overall, lenders are clearly opting to locate in more profitable neighborhoods.

This preference for profitable neighborhoods implies that OLS estimates of the effect of branch presence will be biased toward greater mortgage access and higher borrower quality. For example, in tracts with a growing population of profitable borrowers, lenders should sell fewer loans to the secondary market, experience lower foreclosure rates, lend to more borrowers with higher incomes, and originate more loans as conventional. When lenders locate branches in these tracts, OLS results at any level of aggregation will reflect that bias for each outcome.

As predicted, the OLS results for lender-specific and aggregate mortgage effects do suggest that branches expand mortgage access and shift lending to more profitable borrowers. In Table 4, I show that a lender's own branches and competitors' branches *both* have a positive correlation with a lender's supply of loans to a neighborhood. Furthermore, the sign of the coefficient on percent sold, three-year foreclosure rates, and percent conventional are all in the direction of the expected bias. In Table 5, OLS results for census tracts without a previous branch show a similar pattern and, most importantly, that branches increase aggregate credit access in these tracts. Together, a naive interpretation of these results could be that branches improve credit access for borrowers – perhaps through the effects of competition and the collection of soft information.

First-stage estimates.

Identification of the causal effect of branch presence in this study depends on the relevance of my network density instruments. I find that the density of branches surrounding a tract has a strong and highly statistically significant impact on the presence of a lender's branches in a census tract. Columns 1-4 of Table 6 show that the presence of a lender's own branches in each ring of distance has an independent effect on their own branch presence. And column 5 shows that these coefficients are very stable when measures of distance to the overall network and competitors' branches are

[29] The percent of households with a degree was included as a possible correlate of search costs, assuming a high concentration of households without degrees would indicate an area with less knowledge about the mortgage market. I find almost no correlation, which is somewhat unsurprising given the difficulty of measuring search costs at the neighborhood level. But, it at least appears that lenders do not avoid areas with less educated households, holding other features fixed.

included. The density of branches around a census tract similarly determines competitors' presence, except that competitors prefer locations away from another lender and near their own branches. These differential effects suggest that, despite being predicted by the same set of instruments, the predicted number of own and competitor branches will be well identified.

Instruments based on overall branch and lender density for the presence of any branch in the subsample of tracts before and after their first branch also satisfy the relevance condition. Table 7 shows that the strongest instruments for the first branch are the density of branches around the tract and the number of lenders operating at least 5 branches in each ring of distance from the tract centroid. The number of branch network centers in each ring of distance is not as strong a set of instruments, but given the relevance of very close centers 5-10 miles away, I also include these instruments in my estimation. Column 4 shows the combined significance of my preferred set of instruments and serves as the first stage for estimating the effect of branches in these tracts. Other first stage instruments with similar evidence can be found in Appendix A.

The instrumented branch variables mean that the IV results will rely on different variation in the data to identify causal effects. The lender-specific IV results will be identified off of the differences in a specific lender's mortgage outcomes in tracts that have different numbers of predicted branches for each lender, rather than their actual number of branches in each tract. Similarly, the aggregate IV results will be identified off of the difference in aggregate mortgage outcomes in tracts with different numbers of predicted total branches, rather the actual total number of branches in each tract. These predictions from the instruments correct for the endogeneity of branch location choice and allow for the causal interpretation of the second-stage estimates. The comparison of the OLS and IV estimates for the effect of branches on lender-specific and aggregate lending will show that these instruments are key to capturing the true effect of branches on local mortgage outcomes.

Hypothesis 2: Lender-specific effects.

The IV results for all banks, reported in Table 8, provide strong evidence that lenders respond to the presence of their own and competing branches as if branches provide an informational advantage. The IV results still show that a one-standard deviation increase in a bank's predicted branches in a census tract increases their lending overall and to more profitable borrowers, consistent with the findings from the OLS results. However, the effect of competing bank branches is very different. The IV results show that, rather than having almost no effect (as in the OLS results), competing bank branches increase the denial rates of other lenders and shift their lending toward more conventional borrowers, who must have high-quality hard information to qualify for those loans. These effects are consistent with an adverse selection problem due to informational asymmetries in borrower soft information.

Separating banks into local and non-local banks and branches by their specialty produces results that suggest organizational structure affects the informational advantage conveyed by branches. In particular, branches owned by local banks and specializing in mortgages appear to use soft information to significantly increase lending to low-quality hard-information borrowers. As seen in Panel A of Table 9, a one standard deviation increase in these branches leads to a decrease of 0.9 percentage points for conventional mortgages and a decrease in average income of 0.9 percentage points (about $1,000). Despite the decrease in the observational quality of borrowers this implies, local banks appear to view these loans as more profitable since they decrease the percent sold by 3.1 percentage points.

In contrast, other branches appear to be be used by lenders to primarily cream-skim high-quality hard information borrowers. For example, Panel B of Table 9 shows that for non-local banks, a one standard deviation increase in the presence of either branch type has a large negative impact on the percent of loans sold to the secondary market and foreclosure rates, driven by a shift to more borrowers that qualify for conventional loans. These differential effects of lender size and branch type are consistent with the hypothesis that the effects of soft information are strongest in cases where lenders have the skill and ability to process it.

The patterns in individual bank total mortgage supply and demand provide additional evidence that adverse selection drives this lending behavior. In the last three columns of Table 9, it is clear that branches can be a powerful way to attract and choose borrowers from a larger pool of applicants. Applications, total originations, and denial rates all increase for every own branch type – except for mortgage branches owned by local lenders which show no change. But the more important point is that applications and denial rates also increase with the number of competing branches, regardless of bank size. The matching and searching of borrowers for lenders described in the model seems like the best explanation for this fact; more low-quality applicants could be applying to lenders without branches after receiving less attractive offers from lenders with branches.

Hypothesis 3: Aggregate effects.

Given the potentially offsetting effects of different lenders with and without branches, the expected aggregate effect of a branch on mortgage origination features and credit access is unclear. Empirically, it could be positive or negative. Using the subsample of census tracts before and after they acquire their first branch, I first estimate the instrumented aggregate effect of a bank branch when branches owned by mortgage brokers and non-bank lenders are excluded from the analysis. Table 10 shows no evidence that the adverse selection harms aggregate credit access. In fact, the average income of borrowers decreases by 5.7 percentage points, suggesting that branches do shift lending toward low-quality hard information borrowers. Ending the investigation here would, again, lead to the conclusion that even if adverse selection affects individual lenders, the aggregate effect of branches is to improve access for low-socioeconomic status borrowers.

However, excluding points of mortgage access from the analysis may create measurement error that biases the true effect. Like banks, non-banks lenders and mortgage brokers endogenously choose the location of their branches. For example, if they prefer higher (lower) income neighborhoods, then the miss-measurement of the aggregate number of branches will be higher (lower) in those neighborhoods. The non-negative correlation between income and the size of the error term will then bias the effect of a branch.

In fact, once mortgage broker and non-bank branches are included in the aggregate analysis, I find strong evidence of an adverse selection problem in this subsample. The first column of Table 11 shows that a one standard deviation increase in branches causes a 4.2 percentage point decrease in the total percent of loans sold to the secondary market. This can be explained by lenders' dramatic shift toward high-quality hard information borrowers. In aggregate, average borrower income increased by 10.2 percentage points and the percent originated as conventional increased by 9.0 percentage points. Though the increase in foreclosure rates of 1.9 percentage points suggests that higher quality was not entirely borne out, overall these results suggest that the introduction of a branch into a neighborhood without a previous branch reduces credit access for low socioeconomic status borrowers.

Separating branches by type and lender size again uncovers significant heterogeneity that supports soft information as the force driving these effects. As shown in Table 12, in tracts without a previous branch, general service bank branches shift credit away most strongly from borrowers with low-quality hard information. A one standard deviation increase in general service bank branches increases average borrower income by 31.8 percentage points and increases the percent of loans originated as conventional by 16.6 percentage points. The same problem affects census tracts that receive a bank branch specializing in mortgages and, more significantly, causes a 2.7 percentage point increase in percent sold to the secondary market and 1.8 percentage point decrease in three-year foreclosure rates. The strong effect for these two measures, in particular, suggests that lenders perceive competing mortgage branches to convey such a significant advantage that their mortgages from that neighborhood should be moved out of portfolio to avoid losses. Given that these branches are more likely to have the in-house expertise to collect and use soft information, the soft information is a likely mechanism for changes in lender behavior in response to branches.

Additional support for the view that soft information through branches drives lender behavior is seen in the response to branches owned by local banks. I find that market segmentation is the dominate result, rather than adverse selection, when that general service or mortgage bank branch is owned by a small, local bank.[30] In Table 12, the coefficients on the effect of a local branch move the total effect for a local mortgage or general service bank branch back toward zero. For mortgage branches owned by a local lender, the combined effect is a decrease of 0.3 percentage points for percent sold, indicating greater overall expected profitability. Furthermore, being a local branch reduces the increase in average borrower income to only 7.3 percentage points and the percent originated as conventional to 6.5 percentage points. The less severe adverse selection is consistent with local banks' ability to use soft information to specialize in low-quality hard information borrowers and segment the market rather than compete directly with non-local lenders.

The effect of branches on aggregate mortgage credit availability is dependent on the severity of the adverse selection problem. In the final three columns of Table 12, a one standard deviation

[30] Bank branches are either general service or mortgage specialists and owned by a local or non-local bank. There is not enough variation in the data to measure a separate interaction effect for banks that are of each type and owned by local bank. Branches specializing in mortgages are relatively uncommon.

increase in branches has no detectable effect on log loans, log applications, or percent of originations denied. But when branches are separated by type and lender size in Table 6, the final three columns show that general service bank branches owned by non-local banks actually reduce the overall amount of credit. The reduction in mortgage credit for low-quality hard information borrowers by lenders without a branch appears to dominate any increase in lending by the bank operating the branch. Local bank branches, which showed a less severe adverse selection problem in the first three columns, are shown in the last three columns to increase aggregate credit. In this case, it appears that market segmentation allowed the local bank branch to cater to low-quality hard information borrowers without drastically reducing the mortgage activity of other lenders.

Branches beyond the first branch produce results that are broadly consistent with the effects of the first branch. Specifically, I examine the effect of an additional lender operating at least one branch of each type and size on the same mortgage outcomes. These results can be found in Table 13. The effects of lenders operating general service bank branches, bank branches specializing in mortgages, and the difference of those effects when owned by a local lender are similar, but somewhat diminished from the effects of the first branch. This is expected; the more competitors with branches there are, the more lenders have access to soft information and the less any individual lender has a large competitive advantage.[31]

However, the results for mortgage brokers and non-bank lenders are much more statistically significant in this larger sample. A one standard deviation increase in mortgage brokers operating a branch in a census tract increases the percent of loans sold on the secondary market by 0.6 percentage points. At the same time, the percent of loans originated as conventional increases by 1.9 percentage points and average borrower income increases by 1.9 percentage points. Although it could be argued that this aggregate effect is not working through the increased screening on soft information by mortgage brokers, credit clearly shifts away from borrowers with low-quality hard information. In contrast, non-bank lenders who do not suffer from such agency problems and have incentive to use the available soft information, have a negative effect on the percent of originations sold of 0.8 percentage points. Like banks, the incentive structure of mortgage brokers and non-bank lenders affects the severity of adverse selection, with broker branches increasing aggregate denial

[31]The negative coefficients on log loans and log applications for local lenders in Table 7 is likely the result of the presence of lenders who are not required to report their loans to HMDA.

rates and non-bank lenders decreasing aggregate denial rates.

VI Conclusion

In contrast to previous research, I show that asymmetries in soft information are present in the mortgage market and adversely affect low socioeconomic status borrowers. I build evidence for this by examining changes in lenders' mortgage supply in the presence of their own and competitors' branches and how their combined responses affect aggregate credit outcomes. Branches attract a substantial share of applicants for the operating lender, which allows them to screen for the most profitable borrowers using soft information. Lenders respond to other lenders' branch advantage by raising their own credit standards. These responses are symptomatic of a classic adverse selection problem, such that, in aggregate, I find that a branch reduces the share of credit going to low socioeconomic status borrowers.

These outcomes vary with branch type and lender size in ways that closely correspond with theory and provide additional evidence that soft information is the mechanism affecting lenders' supply decisions. Local lenders adversely affect low-quality hard information borrowers the least, particularly when they specialize in mortgages. Their incentive and ability to cater to borrowers of these types likely allows them to specialize in lending with soft information and segment the market. Local lenders are also found to increase aggregate mortgage credit despite some adverse selection, unlike non-local lenders who create such a strong adverse selection problem that they reduce it.

This paper makes original use of lenders' branch network optimization problem to measure the effect of branch presence on mortgage credit access through soft information. The instruments derived from this problem are both valid and plausibly exogenous. They are valid because lenders have a strong preference for building dense branch networks – placing new branches close to their other branches and the center of their own network – due to economies of density. My instruments are also exogenous if the presence of branches is unrelated to mortgage outcomes in neighborhoods that are somewhat close by. I argue that this is largely true due to the large fixed costs of operating a branch that create slow-changing branch networks. These two features support a causal interpretation of my results.

These findings are based on a more complete dataset than has yet been used to study mortgage credit access. The excellent coverage and granularity allow me to detect fine changes in local mortgage markets, at both the lender-specific and aggregate neighborhood level. Furthermore, the incorporation of new information on mortgage broker and non-bank lender branches reduces measurement error and is critical to my finding that branches create an adverse selection problem.

This study has important implications for mortgage and housing policy. For policy makers, increasing mortgage access for underserved groups has long been a mission statement. This paper does not speak to whether that, in itself, is an appropriate policy goal. But it does make clear that the policy of encouraging lenders to open branches in underserved areas or create other environments with asymmetric information has consequences that can run counter to that goal. However, increases could still come from more symmetric increases in soft information, such as better automated underwriting systems or additional fields in mortgage applications that could "harden" some types of soft information.

VII References

Agarwal Sumit, Brent W. Ambrose, Souphala Chomsisengphet, and Chunlin Liu. "The Role of Soft Information in a Dynamic Contract Setting: Evidence from the Home Equity Credit Market." *Journal of Money, Credit and Banking*, 43.4 (2011):633-655.

Agarwal Sumit and Robert Hauswald. "Distance and Private Information in Lending." *The Review of Financial Studies*, 23:7 (2010):2757-2788.

Allen, Jason, Robert Clark, and Jean-François Houde. "The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry." *The American Economic Review*, 104.10 (2014):3365-3396.

Acharya, Viral V., Jean Imbs, and Jason Sturgess. "Finance and efficiency: do bank branching regulations matter?." *Review of Finance* 15.1 (2011): 135-172.

Avery, Robert B., Patricia E. Beeson, and Mark S. Sniderman. "Neighborhood information and home mortgage lending." *Journal of Urban Economics*, 45.2 (1999): 287-310.

Beck, Thorsten, Ross Levine, and Alexey Levkov. "Big bad banks? The winners and losers from bank deregulation in the United States." *The Journal of Finance* 65.5 (2010): 1637-1667.

Berger, Allen N., John H Leusner, John J. Mingo. "The Efficiency of Bank Branches." *Journal of Monetary Economics*, 40 (1997):141-162.

Berger, Allen N. and Robert DeYoung. "The Effects of Geographic Expansion on Bank Efficiency." *Journal of Financial Services Research*, 19 (2001):163-184.

Berger, Allen N., Leora F. Klapper, and Gregory F. Udell. The Ability of Banks to Lend to Informationally Opaque Small Businesses. *Journal of Banking & Finance*, 25 (2001):2127-2167.

Berger, Allen N. and Gregory F. Udell. Small Business Credit Availability and Relationship Lending: The Importance of Bank Organizational Structure. *The Economic Journal*, 112 (2002):F32-F53.

Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein. Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks. *Journal of Financial Economics*, 76 (2005): 237-269.

Bester, Helmut. "Bargaining, Search Costs, and Equilibrium Price Distributions." *The Review of Economic Studies*, 55.2 (1988):201-214.

Bharath, Sreedhar T., Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan. "Lending Relationships and Loan Contract Terms." *Review of Financial Studies*, 24.4 (2011):1141-1203.

Blackburn, McKinley and Todd Vermilyea. The Role of Information Externalities and Scale

Economies in Home Mortgage Lending Decisions. *Journal of Urban Economics*, 61 (2007):71-85.

Boot, Arnoud W. A. and Anjan V. Thakor. Can Relationship Banking Survive Competition? *The Journal of Finance*, 55.2 (200):679-713.

Bos, J. W. B. and C. J. M. Kool. "Bank Efficiency: The Role of Bank Strategy and Local Market Conditions." *Journal of Banking & Finance*, 30 (2006):1953-1974.

Bostic, Raphael W. and Breck L. Robinson. Do CRA Agreements Influence Lending Patterns? *Real Estate Economics*, 31.1 (2003):23-51.

Broeckner, Thorsten. Credit-Worthiness Tests and Interbank Competition. *Econometrica*, 58.2 (1990):429-452.

Burdett, Kenneth and Kenneth L. Judd. "Equilibrium Price Dispersion." *Econometrica*, 51.4 (1983):955-969.

Calem, Paul S. Mortgage Credit Availability in Low- and Moderate-Income Minority Neighborhoods: Are Information Externalities Critical? *Journal of Real Estate Finance and Economics*, 13 (1996):71-89.

Calem, Paul S. and Leonard I. Nakamura. Branch Banking and the Geography of Bank Pricing. *The Review of Economics and Statistics*, 80.4 (1998):600-610.

Chang, Angela, Shubam Chaudhuri, and Jith Jayartne. "Rational Herding and the Spatial Clustering of Bank Branches: An Empirical Analysis." Federal Reserve Discussion Paper Series No. 9697-24, (1997).

Degryse, Hans and Steven Ongena. "Distance, Lending Relationships, and Competition. *The Journal of Finance*, 60.1 (2005):231-266.

Degryse, Hans and Steven Ongena. "The Impact of Competition on Bank Orientation." *Journal of Financial Intermediation*, 16 (2007):399-424.

DellAriccia, Giovanni and Robert Marquez. Information and Bank Credit Allocation. *Journal of Financial Economics*, 72 (2004):185-214.

DellAriccia, Giovanni, and Robert Marquez. Lending booms and lending standards." *The Journal of Finance 61.5*, (2006): 2511-2546.

Ergungor, Ozgur E. Bank Branch Presence and Access to Credit in Low- to Moderate-Income Neighborhoods. *Journal of Money, Credit, and Banking*, 42.7 (2010).

Essene, Ren S., and William C. Apgar. The 30th anniversary of the CRA: Restructuring the CRA to address the mortgage finance revolution." Revisiting the CRA: Perspectives on the future of the Community Reinvestment Act (2009): 1-18.

Favara, Giovanni and Jean Imbs. "Credit Supply and the Price of Housing." *American Economic Review*, 105.3 (2015):958-992.

Federal Reserve Bank of Minneapolis Community Affairs Report. A Compilation of State Mortgage Broker Laws and Regulations, 1996-2006. Report 2007-2 (2007).

Felici, Roberto and Marcello Pagnini. "Distance, Bank Heterogeneity and Entry in Local Banking Markets." *The Journal of Industrial Economics*,76.3 (2008):500-534.

Ferreira, Fernando and Joseph Gyourko. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009." NBER Working Paper, (2011).

Goolsbee, Austan and Chad Syverson. "How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines." *The Quarterly Journal of Economics*, 123.4 (2008):1611-1633.

Gormley, Todd A. "Costly Information, Entry, and Credit Access." *Journal of Economic Theory*, Forthcoming.

Hauswald, Robert and Robert Marquez. "Competition and Strategic Information Acquisition in Credit Markets." *The Review of Financial Studies*, 19.3 (2006):967-1000.

Hirtle, Beverly. The Impact of Network Size on Bank Branch Performance. *Journal of Banking & Finance*, 31 (2007):3782-3805.

Holmes, Thomas J. The Diffusion of Wal-Mart and Economies of Density. *Econometrica*, 79.1 (2011): 253-302.

Huang, Rocco R. Evaluating the real effect of bank branching deregulation: Comparing contiguous counties across US state borders." *Journal of Financial Economics* 87.3 (2008): 678-705.

Hughes, Joseph P., William Lang, Loretta J. Mester, and Choon-Geol Moon. "Efficient Banking under Interstate Branching." *Journal of Money, Credit and Banking*, 28.4 (1996):1045-1071.

Jayaratne, Jith and Philip E. Strahan. The Finance-Growth Nexus: Evidence from Bank Branch Deregulation. *The Quarterly Journal of Economics*, 111.3 (1996):639-670.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. "Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency." *The Review of Economics and Statistics*, 46 (2014):1-18.

Johnson, Christian and Tara Rice. "Assessing a Decade of Interstate Bank Branching." Federal Reserve Bank of Chicago Working Paper 2007-03, (2007).

Kerr, William R., and Ramana Nanda. "Banking deregulations, financing constraints, and firm entry size." *Journal of the European Economic Association* 8.23 (2010): 582-593.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics*, 125.1 (2010):307-362.

Lang, William W., and Leonard I. Nakamura. "A model of redlining." *Journal of Urban Economics*, 33.2 (1993): 223-234.

Loutskina, Elena. "The Role of Securitization in Bank Liquidity and Funding Management." *Journal of Financial Economics'*, 100 (2011):663-684.

Ludwig, Eugene A., James Kamihachi, and Laura Toh. "The Community Reinvestment Act: Past Successes and Future Opportunities." Revisiting the CRA: Perspectives on the future of the Community Reinvestment Act (2009): 84-104.

McAfee, R. Preston and John McMillan. "Search Mechanisms." *Journal of Economic Theory*, 44 (1988):99-123.

Mian, Atif and Amir Sufi. "House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis." *American Economic Review*, 101 (2011):2132-2156.

Nguyen, Hoai-Luu Q. "Do Bank Branches Still Matter? The Effect of Closings on Local Economics Outcomes." Working Paper, (2014).

Petersen, Mitchell J. and Raghuram G. Rajan. The Benefits of Lending Relationships: Evidence from Small Business Data. *The Journal of Finance*, 49.1 (1994):3-37.

Petersen, Mitchell J. and Raghuram G. Rajan. The Effect of Credit Market Competition on Lending Relationships. *The Quarterly Journal of Economics*, 110.2 (1995):407-443.

Petersen, Mitchell J. and Raghuram G. Rajan. Does Distance Still Matter? The Information Revolution in Small Business Lending. *The Journal of Finance*, 62.6 (2002):2533-2570.

Presbitero, Andrea F. and Alberto Zazzaro. "Competition and Relationship Lending: Friends or Foes?" *Journal of Financial Intermediation*, 20 (2011):387-413.

Scheessele, Randall M. HMDA Coverage of the Mortgage Market. Housing Finance Working Paper Series, Working Paper No. HF-007 (1998).

Schwartz, Alex. Bank Lending to Minority and Low-Income Households and Neighborhoods: Do Community Reinvestment Agreements Make a Difference? *Journal of Urban Affairs*, 20.3 (1998):269-301.

Stein, Jeremy C. Information Production and Capital Allocation: Decentralized versus Hierarchical

Firms. *The Journal of Finance*, 57.5 (2002):1891-1921.

Stiglitz, Joseph E. and Andrew Weiss. Credit Rationing in Markets with Imperfect Information. *American Economic Review*, 71 (1981): 393-410.

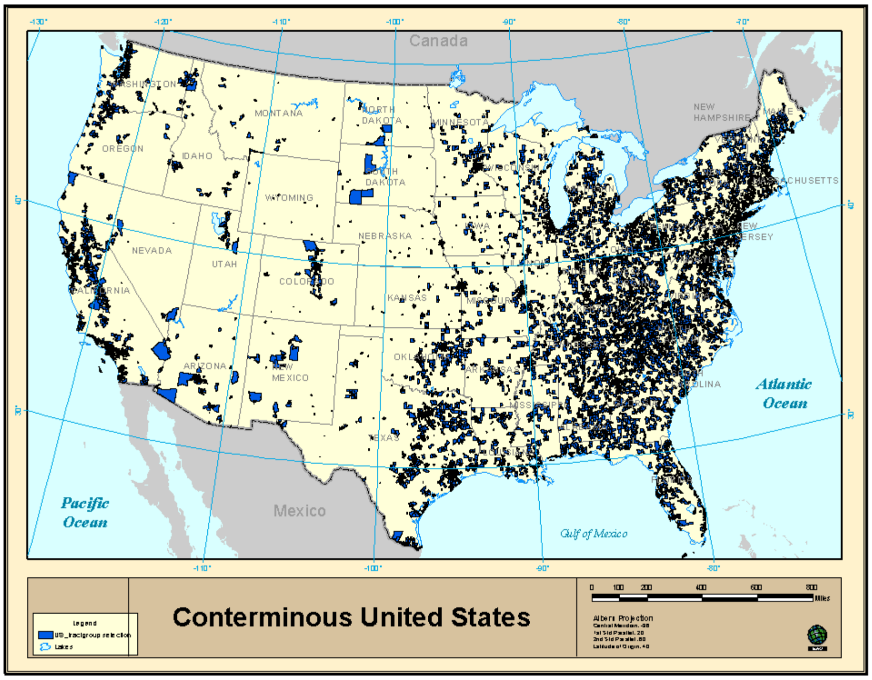
Tewari, Ishani. "The Distributive Impacts of Financial Development: Evidence from Mortgage Markets during US Bank Branch Deregulation." *American Economic Journal: Applied Economics*, 6.4 (2014): 175-196.

Wolinsky, Asher. "Matching, Search, and Bargaining." *Journal of Economic Theory*, 42 (1987):311-333.

Woodward, Susan E. and Robert E. Hall. "Consumer Confusion in the Mortgage Market: Evidence of Less than a Perfectly Transparent and Competitive Market." *American Economic Review: Papers & Proceedings*, 100 (2010):511-515.

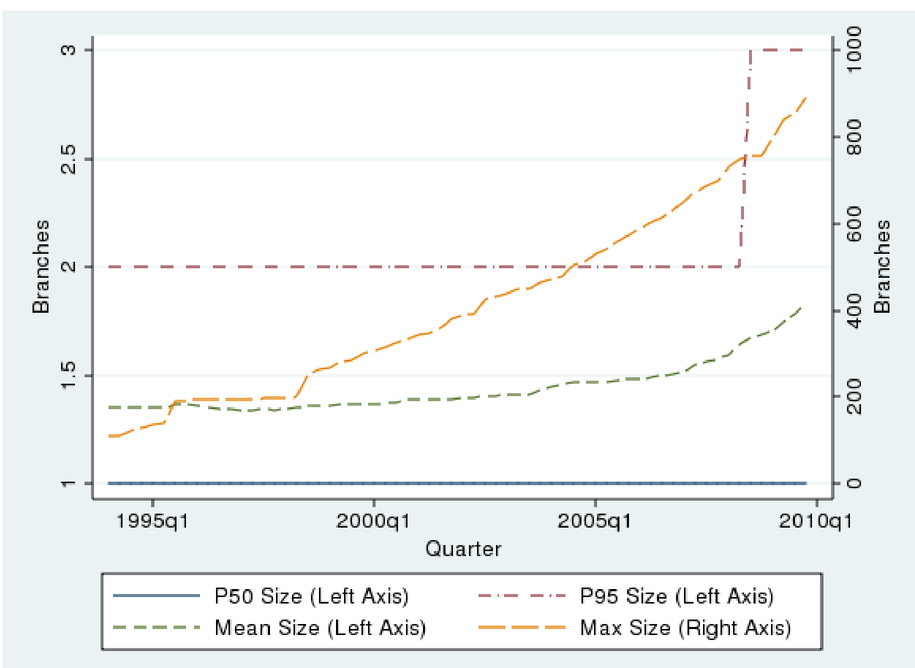
Woodward, Suan E. and Robert E. Hall. "Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence." *The American Economic Review*, 102.7 (2012):3249-3276.

Figure 2: Census Tracts that Receive One Lender Branch



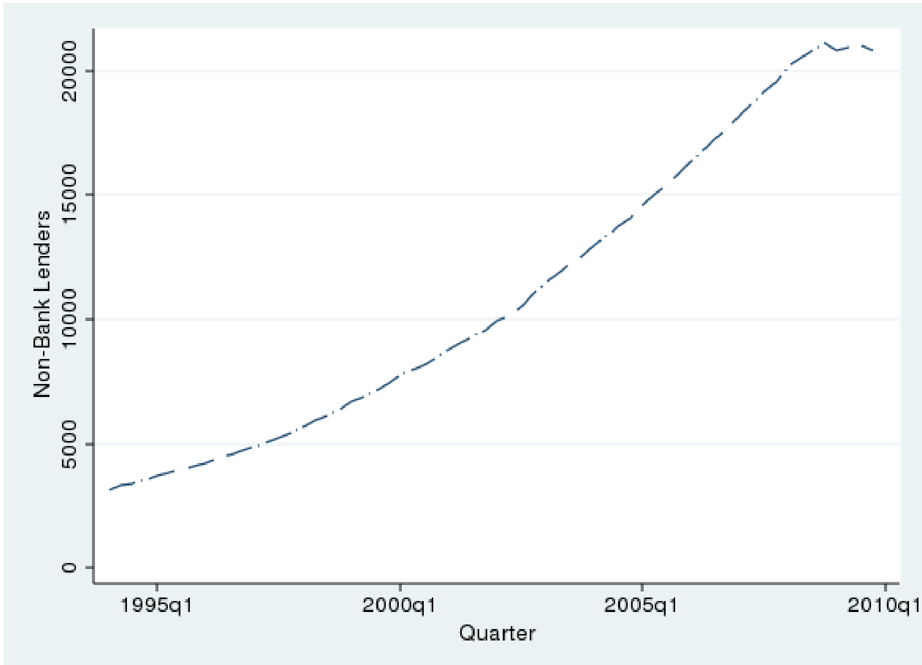
1 Solid blue census tracts are those that receive their first branch during the sample period.

Figure 3: NMLS Lender Branch Network Distribution



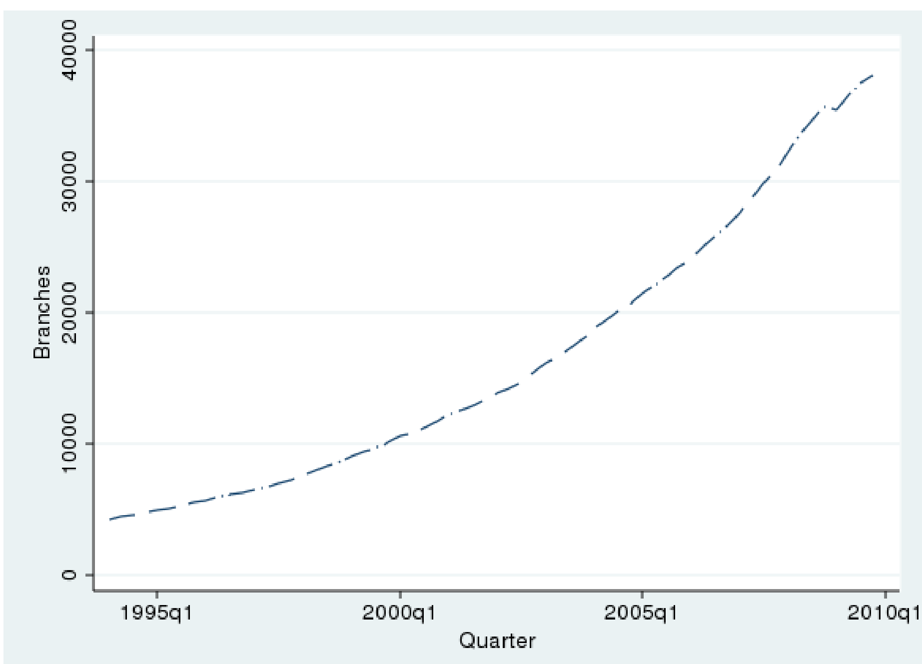
1 Each line is a moment of the NMLS lender branch network size distribtuion. Lenders in this dataset are classified as either non-bank lenders or mortgage brokers.

Figure 4: Number of NMLS Lenders



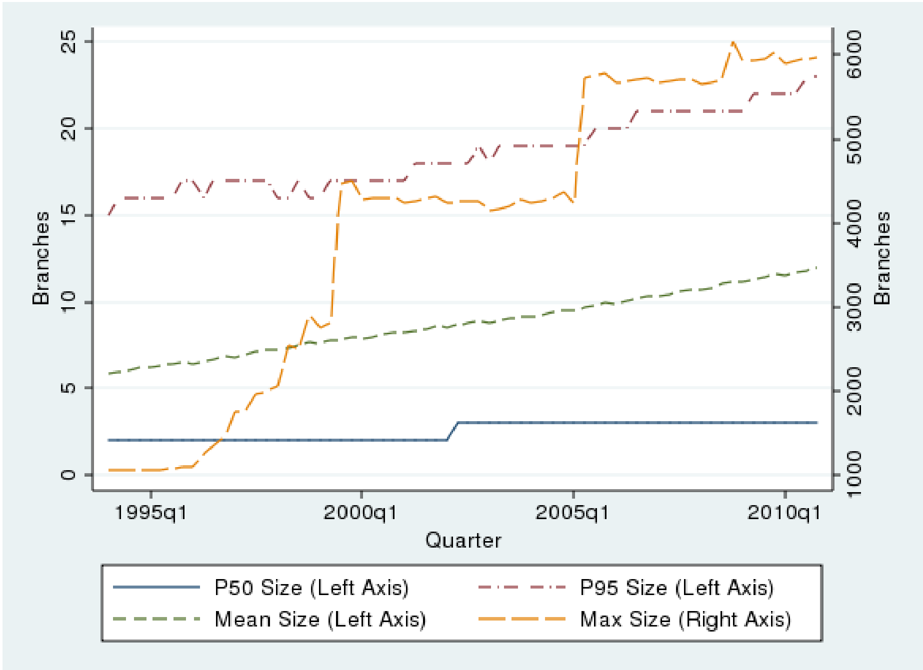
1 The line shows the number of unique lenders in the NMLS data. Lenders in this dataset are classified as either non-bank lenders or mortgage brokers.

Figure 5: Number of NMLS Branches



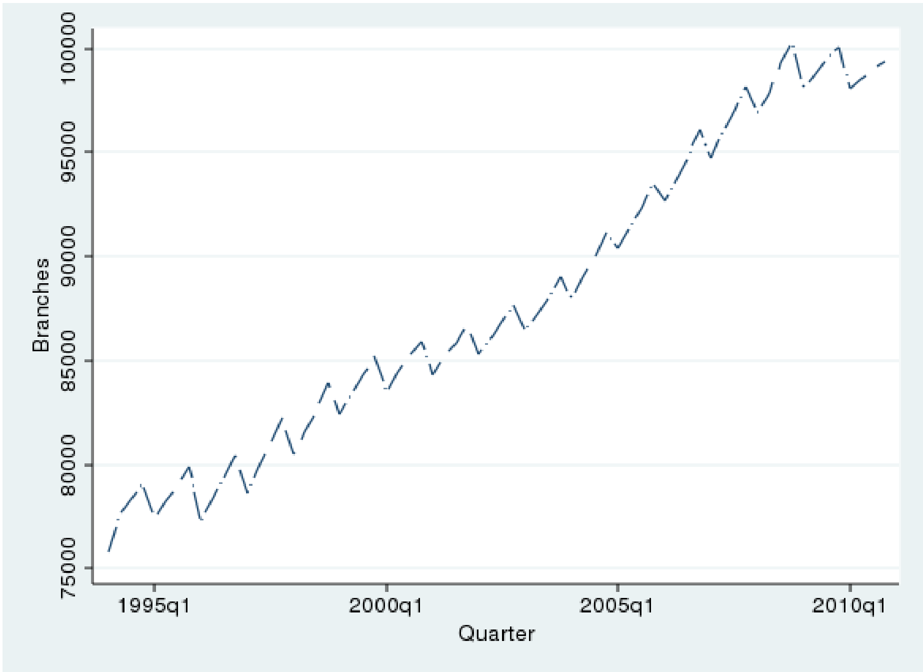
1 The line shows the number of NMLS branches in operation. Branches in this dataset can belong to either lenders or mortgage brokers.

Figure 6: Bank Branch Network Distribution



1 Each line is a moment of the bank branch network size distribtuion. The distribution of network size is extremely skewed, with most networks consisting of a small number of branches. There are a few networks that are much larger.

Figure 7: Number of Bank Branches



1 The line shows the number of bank branches in operation in the FDIC data.

Figure 8: FirstMerit Bank Branch Network, 1996 - 2008.

(a) 1996

(b) 1998

(c) 2000

(d) 2002

(e) 2004

(f) 2006

1 In each year, preexisting branches are shown as small black dots, newly established branches as large blue stars, and acquired branches as large red traingles. In a typical year, the bank adds a few new branches close to its existing branches.

Table 1: Tract Mortgage Characteristics Summary Statistics.

	% Sold		3-Year Fore.		Log Avg. Income		% Conv.	
	Median	SD	Median	SD	Median	SD	Median	SD
1994	0.50	0.22	0.000	0.033	3.93	0.42	0.89	0.18
1995	0.50	0.21	0.000	0.039	3.93	0.42	0.89	0.17
1996	0.50	0.17	0.000	0.043	3.94	0.40	0.90	0.14
1997	0.54	0.16	0.000	0.04	3.98	0.41	0.90	0.13
1998	0.62	0.14	0.000	0.037	4.03	0.38	0.91	0.12
1999	0.57	0.14	0.003	0.037	4.04	0.39	0.91	0.11
2000	0.54	0.14	0.006	0.04	4.08	0.42	0.91	0.11
2001	0.60	0.13	0.004	0.035	4.16	0.40	0.92	0.11
2002	0.64	0.14	0.004	0.036	4.21	0.41	0.93	0.09
2003	0.70	0.14	0.003	0.035	4.25	0.38	0.94	0.09
2004	0.67	0.14	0.006	0.044	4.26	0.40	0.96	0.07
2005	0.68	0.14	0.018	0.056	4.32	0.42	0.97	0.05
2006	0.65	0.13	0.035	0.069	4.39	0.44	0.97	0.05
2007	0.61	0.13	0.030	0.064	4.41	0.45	0.96	0.07
2008	0.66	0.17	0.015	0.052	4.39	0.45	0.82	0.14
2009	0.76	0.18	0.004	0.045	4.37	0.43	0.74	0.19

[1] Table is based on observations at the tract by year level.

Table 2: Tract Mortgage Supply Summary Statistics

	Branches		Log Loans		Log Apps		% Denied	
	Median	SD	Median	SD	Median	SD	Median	SD
1994	1	2	4.42	1.59	4.65	1.52	0.18	0.15
1995	1	2	4.19	1.48	4.48	1.42	0.22	0.16
1996	1	2	4.43	1.19	4.79	1.08	0.27	0.16
1997	1	3	4.52	1.11	4.92	1.01	0.30	0.16
1998	1	3	4.96	1.15	5.34	1.03	0.28	0.16
1999	1	3	4.84	1.08	5.25	0.98	0.32	0.15
2000	1	3	4.61	1.05	5.08	0.96	0.36	0.15
2001	1	3	5.05	1.17	5.42	1.04	0.29	0.15
2002	1	3	5.18	1.22	5.50	1.09	0.25	0.14
2003	1	3	5.46	1.24	5.77	1.12	0.25	0.14
2004	1	3	5.15	1.12	5.53	1.02	0.31	0.13
2005	1	3	5.15	1.13	5.55	1.03	0.32	0.13
2006	1	3	5.06	1.10	5.47	1.02	0.33	0.12
2007	1	4	4.83	1.06	5.29	0.98	0.37	0.13
2008	1	4	4.44	1.08	4.91	0.97	0.36	0.14
2009	1	4	4.56	1.26	4.90	1.11	0.28	0.14

[1] Table is based on observations at the tract by year level.

Table 3: Branch Location Decision

	Branches	Branches	Branches
Households	0.002**	0.006***	0.009***
	(0.001)	(0.001)	(0.001)
Median Income	0.011***	0.009***	0.007***
	(0.001)	(0.001)	(0.001)
%White	0.013***	0.017***	0.012***
	(0.001)	(0.001)	(0.001)
%Degree	0.000	-0.001*	0.000
	(0.001)	(0.001)	(0.001)
%Poverty	0.003***	-0.000	-0.001**
	(0.000)	(0.000)	(0.000)
%Owner Occupied		-0.007***	-0.008***
		(0.002)	(0.002)
%Mortgaged		-0.000	-0.001
		(0.001)	(0.001)
Branches 5-10 M.	0.059**	0.056**	0.056**
	(0.018)	(0.020)	(0.020)
Branches 10-20 M.	0.014*	0.020**	0.020**
	(0.006)	(0.007)	(0.007)
Branches 20-50 M.	0.045***	0.044***	0.045***
	(0.003)	(0.005)	(0.005)
Branches 50-100 M.	0.018***	0.017**	0.018**
	(0.005)	(0.006)	(0.006)
Dist to Center	-0.046***	-0.047***	-0.045***
	(0.006)	(0.006)	(0.006)
Std. Distance	0.040***	0.041***	0.039***
	(0.007)	(0.007)	(0.007)
Comp Branches 5-10 M.			-0.011***
			(0.002)
Comp Branches 10-20 M.			-0.005***
			(0.001)
Comp Branches 20-50 M			-0.017***
			(0.002)
Comp Branches 50-100 M.			-0.020***
			(0.003)
Within R-sq.	0.01	0.01	0.01
No.Obs	583575309	530201529	530201529
Year FE	Y	Y	Y
Lender FE	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county by lender level.

[2] The number of a lender's branches in a census tract is regressed on measures of distance to the lender's own network, measures of distance to competitors' networks, census tract demographics, and census tract housing market statistics. All variables are scaled by their standard deviation.

Table 4: Effect of Bank Branches on Bank's Own Lending, OLS Results

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Branches	-0.005***	-0.001***	-0.001**	0.001***	0.105***	0.111***	0.005***
	(0.001)	(0.000)	(0.001)	(0.000)	(0.002)	(0.002)	(0.000)
Comp. Branches	0.001	0.000	0.001	0.000	0.021***	0.020***	-0.001*
	(0.001)	(0.000)	(0.001)	(0.000)	(0.002)	(0.002)	(0.000)
Within R-sq.	0.05	0.02	0.04	0.07	0.04	0.04	0.01
No. Obs	12117095	4265769	11592318	12117095	12117095	12117095	12117095
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract x Lender FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for $p<0.05$, ** for $p<0.01$, and *** for $p<0.001$. Standard errors are robust and clustered at the county by lender level.

[2] Sample is restricted to observations for banks that have a positive national stock of branches.

[3] Measures of a local and non-local bank's credit supply to a census tract and characteristics of those loans are regressed on the number of that bank's branches of each type and the number of competitor banks' branches in the census tract. All branch variables are scaled by their standard deviation.

Table 5: Effect of the First Branch, OLS Results

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Branch	-0.000	-0.000	-0.000	0.001	0.016***	0.012***	-0.002**
	(0.001)	(0.000)	(0.001)	(0.001)	(0.004)	(0.003)	(0.001)
Within R-sq.	0.27	0.12	0.44	0.34	0.43	0.48	0.20
No. Obs	333985	182198	333366	333985	333985	333985	333985
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for $p<0.05$, ** for $p<0.01$, and *** for $p<0.001$. Standard errors are robust and clustered at the county level.

[2] Sample restricted to census tracts in years when they have 0 or 1 branches.

[3] Measures of the aggregate credit supply to a census tract and characteristics of those loans are regressed on the presence of a branch in the census tract. Branch is scaled by its standard deviation.

Table 6: Instrumenting for Bank Branches

	Branches	Branches	Branches	Branches	Branches	Comp. Branches
Branches 5-10 Mi.	0.123***	0.064***	0.062***	0.062***	0.059***	-0.010***
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.002)
Branches 10-20 Mi.		0.073***	0.020*	0.024**	0.023**	-0.004*
		(0.009)	(0.008)	(0.008)	(0.008)	(0.002)
Branches 20-50 Mi.			0.073***	0.043***	0.041***	-0.001
			(0.005)	(0.005)	(0.005)	(0.002)
Branches 50-100 Mi.				0.042***	0.041***	-0.015***
				(0.005)	(0.005)	(0.001)
Dist to Center					-0.029***	-0.015***
					(0.003)	(0.002)
Std. Distance					0.009**	-0.010***
					(0.003)	(0.002)
Comp. Branches 5-10 Mi.					-0.002	0.067***
					(0.003)	(0.004)
Comp. Branches 10-20 Mi.					-0.022***	0.017***
					(0.003)	(0.003)
Comp. Branches 20-50 Mi.					0.008**	0.066***
					(0.003)	(0.003)
Comp. Branches 50-100 Mi.					-0.009***	-0.031***
					(0.002)	(0.001)
Within R-sq.	0.02	0.02	0.02	0.03	0.03	0.04
No. Obs	12117095	12117095	12117095	12117095	11963211	11963211
Year FE	Y	Y	Y	Y	Y	Y
Tract x Lender FE	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county by lender level.

[2] Sample restricted to observations for banks that have a positive national stock of branches.

[3] The number of a bank's branches in a census tract is regressed on measures of distance to the bank's branch network. These measures include the number of that bank's branches within rings of distance of the tract centroid, the distance to the geographic center of the network, and the standard distance between that bank's branches. Standard distance is a measure of geographic spread between branches. All branch variables, distance to the geographic center, and standard distance are scaled by their standard deviation.

Table 7: Instrumenting for the First Branch

	Branch	Branch	Branch	Branch
Branches 5-10 Mi.	0.008	-0.102	-0.105	-0.132*
	(0.056)	(0.064)	(0.064)	(0.059)
Branches 10-20 Mi.	-0.140**	-0.022	-0.023	-0.161*
	(0.054)	(0.076)	(0.076)	(0.074)
Branches 20-50 Mi.	0.112**	0.180*	0.180*	0.145
	(0.043)	(0.091)	(0.091)	(0.095)
Branches 50-100 Mi.	-0.051	-0.037	-0.041	-0.158*
	(0.039)	(0.080)	(0.079)	(0.069)
5 Branch Lenders 5-10 Mi.	0.037			0.063*
	(0.026)			(0.026)
5 Branch Lenders 10-20 Mi.	0.144***			0.159***
	(0.029)			(0.028)
5 Branch Lenders 20-50 Mi.	0.104**			0.111**
	(0.036)			(0.040)
5 Branch Lenders 50-100 Mi.	0.172***			0.185***
	(0.041)			(0.041)
Branch Centers 5-10 Mi.		0.096	0.097	0.116**
		(0.050)	(0.050)	(0.038)
Branch Centers 10-20 Mi.		-0.057	-0.056	-0.011
		(0.046)	(0.046)	(0.047)
Branch Centers 20-50 Mi.		-0.021	-0.021	-0.036
		(0.049)	(0.049)	(0.052)
Branch Centers 50-100 Mi.		-0.011	-0.008	0.062
		(0.047)	(0.047)	(0.035)
Mean Std. Distance 5-10 Mi			-0.002	
			(0.003)	
Mean Std. Distance 10-20 Mi			0.000	
			(0.004)	
Mean Std. Distance 20-50 Mi			0.004	
			(0.005)	
Mean Std. Distance 50-100 Mi			0.005	
			(0.005)	
Within R-sq.	0.11	0.11	0.11	0.11
No. Obs	333979	333979	333979	333979
Year FE	Y	Y	Y	Y
Tract FE	Y	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county level.

[2] Sample restricted to census tracts in years when they have 0 or 1 branches.

[3] The presence of a branch is regressed on the number of branches within rings of distance of the tract centroid, the number of lenders with at least 5 branches operating in each ring of distance from the tract centroid, the number of branch network centers within each ring of distance from the tract centroid, and the mean standard distance of lenders operating a least 1 branch within each ring of distance from the tract centroid.

Table 8: Effect of Bank Branches on a Bank's Own Lending

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Branches	-0.020***	-0.003***	0.002	0.002*	0.168***	0.206***	0.027***
	(0.005)	(0.000)	(0.003)	(0.001)	(0.012)	(0.012)	(0.002)
Comp. Branches	-0.012*	-0.002**	0.006	0.010***	0.010	0.026	0.011***
	(0.006)	(0.001)	(0.003)	(0.002)	(0.013)	(0.013)	(0.002)
First stage F-stat	169.43	133.36	163.59	169.43	169.43	169.43	169.43
No. Obs	11963211	4218784	11444104	11963211	11963211	11963211	11963211
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract x Lender FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for $p<0.05$, ** for $p<0.01$, and *** for $p<0.001$. Standard errors are robust and clustered at the county by lender level.

[2] Sample restricted to banks that have a positive national stock of branches. Only banks can be matched to individual loan originations.

[3] Measures of a bank's credit supply to a census tract and characteristics of those loans are regressed on the number of that bank's branches and the number of competitor banks' branches in the census tract. The number of the bank's branches is instrumented with measures of distance to that bank's network of bank branches and the number of competitor banks' branches is instrumented with measures of distance to competitor banks' networks of branches. All branch variables are scaled by their standard deviation.

Table 9: Effect of Type of Bank and Branches on a Bank's Own Lending

Panel A: Local

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Gen. Branches	-0.008	-0.008***	-0.007	0.020***	0.361***	0.434***	0.051***
	(0.013)	00.002)	(0.009)	(0.004)	(0.040)	(0.045)	(0.006)
Mort. Branches	-0.031***	0.001	-0.019***	-0.009***	0.013	0.014	0.003
	(0.005)	(0.002)	(0.006)	(0.002)	(0.030)	(0.025)	(0.004)
Comp. Branches	-0.001	-0.003**	0.009*	0.008***	0.022	0.033*	0.007**
	(0.007)	(0.001)	(0.004)	(0.002)	(0.013)	(0.014)	(0.002)
First stage F-stat	28.11	14.80	24.51	28.11	28.11	28.11	28.11
No. Obs	7009411	2356763	6646306	7009411	7009411	7009411	7009411
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract x Lender FE	Y	Y	Y	Y	Y	Y	Y

Panel B: Non-local

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Gen. Branches	-0.022***	-0.002***	-0.002	0.004***	0.187***	0.228***	0.030***
	(0.006)	(0.000)	(0.003)	(0.001)	(0.011)	(0.012)	(0.002)
Mort. Branches	-0.010***	-0.001**	-0.002	0.003*	0.091***	0.111***	0.014**
	(0.002)	(0.000)	(0.003)	(0.001)	(0.017)	(0.012)	(0.004)
Comp. Branches	0.022**	-0.001	-0.018***	0.004*	0.095***	0.128***	0.023***
	(0.008)	(0.001)	(0.005)	(0.002)	(0.024)	(0.024)	(0.004)
First stage F-stat	58.33	42.92	57.21	58.33	58.33	58.33	58.33
No. Obs	4953800	1862021	4797798	4953800	4953800	4953800	4953800
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract x Lender FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county by lender level.

[2] Sample restricted to observations for local banks that have a positive national stock of branches.

[3] Measures of a local and non-local bank's credit supply to a census tract and characteristics of those loans are regressed on the number of that bank's branches of each type and the number of competitor banks' branches in the census tract. The number of the bank's branches of each type is instrumented with measures of distance to that bank's network of bank branches of that type and the number of competitor bank's branches is instrumented with measures of distance to competitor banks' networks of branches. All branch variables are scaled by their standard deviation.

Table 10: Effect of the First Branch, NMLS data excluded

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Branch	0.014	0.004	-0.055**	-0.022	0.023	0.065	0.030***
	(0.008)	(0.005)	(0.020)	(0.021)	(0.041)	(0.038)	(0.006)
First stage F-stat	5.34	5.73	5.33	5.34	5.34	5.34	5.34
No. Obs	333979	182198	333360	333979	333979	333979	333979
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county level.

[1] Sample restricted to census tracts in years when they have 0 or 1 branches of any type. Data on NMLS brokers and non-bank lenders then excluded.

[2] Measures of the aggregate credit supply to a census tract and characteristics of those loans are regressed on the presence of a bank branch in the census tract. The presence of a bank branch is instrumented with the number of bank branches in rings of distance from the tract centroid, the number of banks operating at least 5 branches in each ring of distance from the tract centroid, and the number of geographic centers of bank branch networks in each ring of distance from the tract centroid. Bank branch is scaled by its standard deviation.

Table 11: Effect of the First Branch

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Branch	-0.042***	0.019*	0.097**	0.090***	0.143	0.090	-0.025
	(0.011)	(0.008)	(0.036)	(0.020)	(0.082)	(0.064)	(0.013)
First stage F-stat	8.30	4.95	8.26	8.30	8.30	8.30	8.30
No. Obs	333979	182198	333360	333979	333979	333979	333979
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for $p<0.05$, ** for $p<0.01$, and *** for $p<0.001$. Standard errors are robust and clustered at the county level.

[2] Sample restricted to census tracts in years when they have 0 or 1 branches.

[3] Measures of the aggregate credit supply to a census tract and characteristics of those loans are regressed on the presence of a branch in the census tract. The presence of a branch is instrumented with the number of branches in rings of distance from the tract centroid, the number of lenders operating at least 5 branches in each ring of distance from the tract centroid, and the number of geographic centers of branch networks in each ring of distance from the tract centroid. Branch is scaled by its standard deviation.

Table 12: Effect of the First Branch by Type

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Gen. Branch	0.014	0.017	0.276***	0.166***	-0.309*	-0.165	0.074***
	(0.016)	(0.011)	(0.076)	(0.034)	(0.127)	(0.117)	(0.022)
Mort. Branch	0.027***	-0.018**	0.032	0.031*	0.023	0.083	0.031**
	(0.006)	(0.006)	(0.021)	(0.014)	(0.059)	(0.047)	(0.010)
Local Branch	-0.030**	0.021	-0.206***	-0.101***	0.379***	0.248**	-0.059***
	(0.011)	(0.012)	(0.061)	(0.027)	(0.084)	(0.076)	(0.015)
Broker Branch	0.005	0.001	0.024*	0.017***	0.054	0.059**	0.008
	(0.003)	(0.003)	(0.009)	(0.005)	(0.029)	(0.022)	(0.005)
NBL Branch	-0.008	-0.001	0.020	0.033	-0.008***	-0.032	-0.016
	(0.005)	(0.004)	(0.014)	(0.009)	(0.039)	(0.027)	(0.009)
First stage F-stat	16.27	9.83	15.97	16.27	16.27	16.27	16.27
No. Obs	333979	182198	333360	333979	333979	333979	333979
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county level.

[2] Sample restricted to census tracts in years when they have 0 or 1 branches.

[3] Mortgage outcomes are regressed on the presence of a branch of each type and lender size in the census tract. The presence of a branch is instrumented with the number of branches in rings of distance from the tract centroid, the number of lenders operating at least 2 branches of that type in each ring of distance from the tract centroid, and the number of geographic centers of branch networks in each ring of distance from the tract centroid. Each branch variable is scaled by it's standard deviation.

Table 13: Effect of Lenders with Branches by Type

	% Sold	3-year Fore.	Log Avg Income	% Conv.	Log Loans	Log Apps	% Denied
Gen. Branch Lenders	0.006	0.010	-0.009	0.038*	0.109	0.181	0.054***
	(0.010)	(0.008)	(0.026)	(0.016)	(0.111)	(0.094)	(0.013)
Mort. Branch Lenders	0.002	-0.020**	0.031	-0.001	0.173	0.220**	0.027
	(0.009)	(0.007)	(0.034)	(0.014)	(0.094)	(0.069)	(0.018)
Local Branch Lenders	-0.030**	0.024*	0.030	-0.023	-0.350**	-0.434***	-0.059**
	(0.012)	(0.011)	(0.025)	(0.015)	(0.116)	(0.095)	(0.019)
Broker Branch Lenders	0.006**	0.000	0.019**	0.019***	-0.022	-0.001	0.015***
	(0.002)	(0.002)	(0.006)	(0.004)	(0.022)	(0.018)	(0.004)
NBL Branch Lenders	-0.008**	0.002	0.012	0.008	-0.010	-0.029*	-0.012***
	(0.003)	(0.002)	(0.008)	(0.006)	(0.015)	(0.012)	(0.003)
First stage F-stat	144.31	56.78	144.25	144.31	144.31	144.31	144.31
No. Obs	811427	415565	810384	811427	811427	811427	811427
Year FE	Y	Y	Y	Y	Y	Y	Y
Tract FE	Y	Y	Y	Y	Y	Y	Y

[1] Significance levels are * for p<0.05, ** for p<0.01, and *** for p<0.001. Standard errors are robust and clustered at the county level.

[2] Sample restricted to census tracts in years when they have less than 10 branches.

[3] Measures of the aggregate credit supply to a census tract and characteristics of those loans are regressed on the number of a lenders operating a branch in the census tract. The number of lenders operating a type of branch is instrumented with the number of branches in rings of distance from the tract centroid, the number of lenders operating at least 2 branches of that type in each ring of distance from the tract centroid, and the number of geographic centers of branch networks in each ring of distance from the tract centroid. Each lender coefficient in the second stage is scaled by the standard deviation of its predicted value from the first stage.