Economía, XXXV, 30 (julio-diciembre, 2010), pp. 11-55
ISSN 1315-2467, Depósito legal pp: 198702me336

Branch location choice: Do lenders discriminate?

Elección de localización de sucursales bancarias: ¿Discriminan los prestamistas?

AKM Rezaul Hossain*

Códigos JEL: R2, R3, G2, J15

Recibido: 06/04/2011, Revisado: 25/04/2011, Aceptado: 20/05/2011

Abstract

This paper examines the factors that affect branch location choice of depository lenders. Consistent with previous studies, the tract level analysis in this paper finds a strong negative impact of neighborhood income and minority composition on individual access to depository branches even after controlling for neighborhood unobservables. In the lender level analysis the negative impact of income and race does not persist once neighborhood fixed effects are included. This raises an important incongruity between individual (lender) and aggregate (tract) level analysis. The paper suggests that this inconsistency is related to the logic of disparate treatment and disparate impact by showing how a neighborhood with low income and high minority composition may end up with fewer branches (a disparate impact) even when no individual lender makes branch location choice based on income or race (no disparate treatment).

Key words: Branch location choice, neighborhood effects, lender heterogeneity, disparate treatment and disparate impact of branch location choice.

Resumen

Este artículo examina los factores que afectan la escogencia de la ubicación de las sucursales depositantes prestamistas. Consistente con estudios previos, el análisis del nivel de información agregada geográfica en este artículo revela un fuerte impacto negativo del nivel del ingreso del vecindario y de la composición minoritaria sobre el acceso individual a las sucursales bancarias prestamistas incluso después de ser controladas por el vecindario inobservable. En el nivel de análisis de sucursales bancarias prestamistas, el impacto negativo del ingreso y la raza no persisten una vez que el los efectos fijados por el vecindario se incluyen. Esto produce una importante incongruencia entre el nivel

* Division of Business, Mont Saint Mary College (MSMC), Newburgh, New York. Correo electrónico: hossain@msmc.edu.

del análisis individual (del prestamista) y el nivel del análisis agregado (de información agregada geográfica). El artículo sugiere que esta inconsistencia está relacionada con la lógica del tratamiento y del impacto esencialmente diferenciados al mostrar cómo un vecindario con bajo ingreso y alta composición minoritaria puede terminar con pocas sucursales (un impacto diferenciado) incluso cuando ninguna prestamista individual elige la ubicación de una sucursal basada en el ingreso o en la raza (tratamiento no diferenciado).

Palabras clave: Elección sucursal bancaria, efectos locales, heterogeneidad de los prestamistas, tratamiento diferenciado, impacto diferenciado.

1. Introduction

The availability of banks in a neighborhood is important for its residents and local businesses. Some argue that the flow of credit and financial services accorded by banks is a key element for economic growth and vitality of communities.[1] Others argue that familiarity with basic banking services allows people to save for the future and opens up the broader credit market, which is crucial for upward mobility in society.[2] However, the distribution of bank branches is often highly clustered potentially resulting in different access to credit and financial services across neighborhoods.[3] Since location is the primary determinant of where people conduct banking activities,[4] lack of traditional banks in a neighborhood may leave its residents unbanked or force them to fulfill their banking needs at fringe banks such as check cashers, payday lenders and pawn shops. Given the immense significance of credit and financial services in personal and economic life, it is important to know why certain neighborhoods lack bank representation.

Depository institutions contribute to economic growth by specializing in allocating financial resources from savers to borrowers at a lower transaction and information cost. In providing this service, they make a profit from the net interest margin, which is the difference between interest earnings from lending and interest expense for attracting deposits. According to Felici and Pagnini (2004), this profit objective plays an important role in individual lender's decision to enter into a local banking market. Although this profit seeking behavior is at the heart of efficient allocation of financial resources, researchers show that neighborhood median income and racial composition significantly explains the aggregate distribution of branches across neighborhoods. Specifically, they find that poor and minority dominated neighborhoods are more likely to have below average branch representation. Given this aggregate outcome, this paper attempts to understand if individual lenders stay away from poor and minority neighborhoods when they make branch location choice.

This paper estimates models of branch location choice of depository lenders using a rich panel that combines detailed branch location information with lender and neighborhood characteristics for the state of Connecticut. Unlike previous studies[5] that use a cross section of branch location information from multiple states and cities, the panel nature of this dataset allows for fixed effects to capture time-invariant, unobserved neighborhood and lender heterogeneities associated with branch location choice of depository lenders. Although the importance of heterogeneity has been generally recognized in previous studies,[6] this paper estimates models of lender specific branch location choice with neighborhood fixed effects for the first time.

Consistent with the previous studies, the tract level analysis in this paper finds that low median income and higher minority concentration significantly reduce aggregate branch representation by all lenders even after controlling for unobserved neighborhood attributes. However, in the lender level analysis this relationship does not persist once neighborhood fixed effects are controlled for suggesting that neighborhood median income or minority concentration have no role in branch location choice of individual lenders. This raises an important fallacy of composition in which although individual lenders don't stay away from poor and minority neighborhoods, aggregate behavior of all lenders appears to suggest just that. The paper proposes that this apparent incongruity between individual and aggregate behavior can be can be interpreted as the market having a disparate impact on low income and minority neighborhoods. This relates to the arguments made by Ross and Yinger about how disparate impact discrimination may arise in mortgage lending in an aggregate model while little or no difference may be detected in the lender specific models (Ross and Yinger, 2002).

In making branch location choice, if a lender deliberately stays away from low income and minority dominated neighborhoods this action can be considered as disparate treatment discrimination or redlining against those neighborhoods. However, whenever certain neighborhoods observe disproportionately low representation of branches even though no adverse treatment is practiced we may regard this outcome as disparate impact upon those neighborhoods.[7] The paper shows that disparate impact upon low income and minority dominated neighborhoods results from unobserved neighborhood attributes that fulfill two properties simultaneously. First, these neighborhood heterogeneities are correlated with neighborhood income and racial composition. Second, they affect individual lenders' branch location choice. The paper argues that only a lender level analysis with neighborhood fixed effects can distinguish disparate treatment from disparate impact of branch location choice. The traditional tract level analyses are inadequate in this respect.

The remainder of the paper is organized in six sections. Section two provides an overview of previous studies and their findings. Section three describes the empirical model and methodology. Section four discusses the data and construction of several key variables. Section five describes the results and six concludes the paper.

2. Overview of branch location studies

Little is known about the branch location choice of depository lenders. Two papers directly focus on branch representation and several other related papers provide additional insights[8] on the topic. Using branch location addresses from telephone book yellow pages, Avery (1991) compared the size and distribution of branch offices of banks and other financial institutions[9] in 1977 and 1989 in five metropolitan areas[10] to understand the impact of branching deregulation that has taken place over the period of that study. Compared to 1977, bank branches per capita have increased in all zip codes except in the zip codes with high concentrations of Black residents. The paper finds that zip code areas with low median income[11] or with high concentration of Black residents[12] had significantly fewer branches per capita in 1989 compared to areas with predominantly high-income or predominantly white residents for all five metropolitan areas. However, when other relevant demographic factors[13] are included, the impact of race and income on per capita number of branches remains negative but its significance disappears except for the effect of race on thrift institutions.

Using a probit model of branch location data for five major cities[14] from 1970 to 1989, Caskey (1994) finds that the communities with majority of Black residents are significantly less likely to have a bank branch compared to other communities in all five cities. In Atlanta and New York City, both low-income and minority communities are less likely to have a bank branch. In Denver and Washington DC, a similar result is found only for minority communities. In San Jose, neither low-income nor minority communities are less likely to have a branch compared to high income, white communities. Caskey (1994) refers to two earlier papers[15] that find lower per capita representation of financial institutions in areas with high percentage of low-income and minority population.

Using data from 1990 through 1995, Chang *et al*. (1997) empirically explores whether spatially clustered bank branches in New York City are attributable to rational herding[16] by banks. Using Poisson and ordered logit estimation, their paper finds that after controlling for key neighborhood characteristics,[17] the number of preexisting branches in a tract positively and significantly affects the number of new branch openings in the tract implying evidence of herd behavior in branch opening decision. Their paper also finds a significant negative impact of the proportion of poor and minority population in tract on branch openings. The authors point out that this finding may be due to unobserved profitability factors correlated with neighborhood income and racial composition. Due to data constraints, their paper could not implement a tract fixed effects model, which would have been appropriate to control for unobserved neighborhood characteristics.[18] The authors, however, pursue an alternative approach to test for the severity of omitted variable bias. Specifically, they regress expected profitability on number of branch openings postulating that absent herding behavior branch openings must be motivated by profitability factor. However, they find that branch openings in New York City tracts are negatively related to profits, which led the authors to suggest that branch opening decisions in their study do not reflect unmeasured profitability. Rather, these decisions are attributable to herding behavior.

The regulations arising from the Community Reinvestment Act (CRA) may have significant implications for depository lenders' branch location choice. The CRA requires depository institutions to help meet the credit needs of the entire community, including low- and moderate-income neighborhoods, in a manner consistent with safe and sound operation.[19] The Act requires lenders to define a specific assessment area where each lender commits to focus its CRA related efforts.[20] The assessment area, however, cannot arbitrarily exclude surrounding low- and moderate-income areas. During periodic CRA examinations, regulators carefully examine the reasonableness of assessment areas. Based on CRA performance, lenders receive a rating that plays an important role in obtaining regulatory approval for future merger, acquisition, and opening and closing of branches. Consequently, a poor CRA performer may have difficulty in expanding, reducing or relocating its branch network. For example, in 1989 the merger application of Continental Bank Corporation to acquire Grand Canyon Bank of Scottsdale was denied on CRA ground.[21] Avery *et al.* (1997) note that to achieve a good CRA compliance record, an institution may open or retain offices in lower income communities.

Antonakes (2001) analyzes the effectiveness of CRA by determining whether individual bank's decision to choose an assessment area leads to sufficient banking services in all communities. Using 1995 data for the state of Massachusetts, this paper finds that the percentage of minority population in a neighborhood negatively and significantly affects number of banks that include the neighborhood in their assessment area. This paper finds opposite impacts for tract median income and population density. Tracts with higher median incomes and population densities significantly increase the number of banks that include these tracts in their assessment area.

Avery (1991), Caskey (1994), Chang *et al*. (1997) and Antonakes (2001) studies find a significant effect of income and minority composition of neighborhoods on bank representation. These studies also recognize the importance of unobserved neighborhood heterogeneities in influencing branch representation across neighborhoods. In that, lenders may have private information[22] about neighborhood characteristics that are not available to researchers. These characteristics may have significant impact on branch location choice of lenders. However, omitting these characteristics from regression equation would bias the coefficient estimates of the included variables in an upward (downward) direction when omitted characteristics are positively (negatively) correlated with the included variables. This paper addresses this issue and controls for unobserved neighborhood characteristics that remains unchanged between years.

In addition to these aforementioned studies, several other papers point out several key neighborhood and lender attributes that may influence lenders' location choice. Cohen and Mazzeo (2004) find that market concentration and lender type[23] significantly affect the quality of products offered by the lender. In their paper, the product quality is the extensiveness of branch network measured by number of branches. Felici and Pagnini (2004) points out the importance of several factors including larger market size, size of potential entrants, profitability of the entrants and pre-existing location of the entrant in the entry decision into local banking markets. In addition, using survey data on households' branch switching behavior at depository institutions, Kiser (2002) reports that from customers' perspectives the most important factor of choosing a branch is the location of the branch relative to the consumer's residence or work place and a low price elasticity of switching depository institutions.[24] This implies that communities with growing populations and emerging businesses will attract potential entrants.

Previous studies on branch location choice can be broadly categorized into two sets. One set of studies focuses on the aggregate changes in number of branches analogous to the tract level analysis of this paper. This set includes studies by Avery (1991), Caskey (1994), Chang *et al*. (1997) and Antonakes (2001). The other set focus on estimating lender specific models including such studies as Mazzeo (2004), Felici and Pagnini (2004) and Kiser (2002). This paper for the first time estimate lender specific models with tract fixed effects. Lender decisions to locate close to a neighborhood can be broadly thought of as a lender's willingness to provide banking and financial services[25] to the residents and businesses of that neighborhood. Although physical presence is an important aspect of providing banking services, it leaves out other crucial quantitative and qualitative dimensions of banking services. For example, hours of operation, number and expertise of staff, availability of drive up ATMs, and the number and availability of banking-related services[26] are important aspect of service not reflected in the number of branch variable. Therefore, it is important to keep in mind while interpreting results presented in the paper that certain lenders may serve the community well without an extensive branch network and certain lenders may have a physical presence with minimal services.

3. Model and methodology

Methodology employed in this paper has four distinct features, which are described in this section under four headings. Those are (*a*) Two levels of analysis –aggregate tract and lender level, (*b*) OLS and Poisson estimation, (*c*) Neighborhood and lender fixed effects and (*d*) Use of census and HMDA based neighborhood characteristics.

3.1. Two levels of analysis

To estimate the impact of factors that influence lenders' decision to locate in a neighborhood, I set up models with two levels of analysis: tract-level analysis and lender-level analysis. In the tract-level analysis, each tract is a unit of observation. Therefore, in this analysis, we estimate the aggregate location choice of all lenders in a tract. I hypothesize that this aggregate location choice can be explained by several tract level socioeconomic characteristics. This can be expressed by the following statistical model:

$$NOB_{ty} = \beta_0 + \beta_1 N_t + \varepsilon_{ty}$$

Here, the dependent variable, NOB_{ty} is total number of branches of all lenders that are within one-mile distance from the center of tract *t* in year *y*. This variable is intended to capture local presence of depository lenders.[27] The variation in the dependent variable would be explained by several neighborhood characteristics, N_t. Specifically, nine neighborhood characteristics are included. Those are tract median income, racial composition (Percentage Black,[28] Percentage Asian, Percentage Hispanic), gender composition (Percentage Male[29]), Percentage Old,[30] vacancy rate (Percentage Occupied), tenure choice (Percentage Owner Occupied) and Population Density. The unexplained residual (ε_{ty}) is assumed to be independent and identically distributed. The neighborhood characteristics in census data do not change between years. However, the eight year pooled data is corrected for robust clustered standard error for all estimations.

The baseline lender level analysis is set up by replacing the dependent variable of the tract level analysis. For this analysis, the dependent variable is modified to a lender-tract specific variable (NOB_{lty}) that counts the number of branch of lender *l* that are within one-mile from the center of tract *t* in the year *y*.[31] Initially, we estimate the lender level analysis using the same control variables used in the tract level analysis. This estimation will allow us to compare the effects of neighborhood characteristics on aggregate branch distribution in the neighborhood by all lenders and on branch location choice of individual lenders.[32] The empirical model for lender level analysis can be expressed as:

$$NOB_{lty} = \beta_0 + \beta_1 N_t + \varepsilon_{lty}$$

Research suggests that lender heterogeneity matters in entry decision to local banking markets.[33] The lender level analysis is extended by including several lender-tract-year (L_{lty}) and lender-year (L_{ly}) controls that are expected to influence number branches of the individual lender across neighborhoods. This analysis is estimated using the following statistical model:

$$NOB_{lty} = \beta_0 + \beta_1 N_t + \beta_2 L_{lty} + \beta_3 L_{ly} + \beta_4 D_l + \varepsilon_{lty}$$

Here, I include four additional lender specific controls: lender's head office location from tract center (Head Distance), dollar amount of deposits received by the lender (Deposit), dollar amount of equities held by the lender (Equity) and average of three past CRA ratings (Past CRA Rating). The Head Distance variable calculates the distance between the center of tract *t* and head office location of the lender *l* that has a branch within a one mile radius of the tract center in year *y*. Therefore, the Head Distance is a lender-tract-year specific variable (L_{lty}). However, Equity and Past CRA Rating are only lender-year specific (L_{ly}) variables that vary between lenders and years, but remain unchanged between tracts.

3.2. OLS and Poisson estimation

The tract and lender level analyses are estimated using both Ordinary Least Square (OLS) and Poisson regression in all specifications. Since the dependent variables for these analyses (NOB_{ty} and NOB_{lty}) take non-negative values with large number of zeros, a Poisson error distribution is assumed. In particular, for the tract level analysis the dependent variable NOB_{ty} or the number of branches of all depository lenders in tract t in year y is assumed to have a Poisson distribution with parameter λ_{ty}, where λ_{ty} is a log linear function of the exogenous variables X_{ty} that represents nine neighborhood characteristics.

$$\ln\lambda_{\mathrm{ty}} = \delta_t + \beta X_{ty}$$

To model branch distribution across neighborhoods similar to the one presented in the baseline tract level analysis of this paper, Caskey (1994) also assumes a Poisson error distribution. Chang *et al*. (1997), however, assume an Ordered Logit distribution for the dependent variable that counts number of new branch openings. I preferred a Poisson over Ordered Logit distribution because number of branches in a tract in any given year can take any non-negative integer, which may not fall into a pre-ordered group.

3.3. Neighborhood and lender fixed effects

Research finds that neighborhood characteristics are important determinants of lending decision and affect the probability of mortgage approvals.[34] It is reasonable to assume that neighborhood characteristics will have an important influence on lenders' decision to locate close to the neighborhoods that help improve their lending prospects. However, not all neighborhood characteristics affecting lenders' location choice are observable. For example, I have no information on the lending risk associated with neighborhood population or businesses. Similarly, no information is available on the degree to which local governments welcome businesses into their neighborhoods. Leaving out these important factors from the estimating equation would bias the coefficient estimates. Similar bias and inconsistency will be generated for unobserved and uncontrolled lender attributes in the lender level analysis. One of the ways to control for time-invariant unobserved neighborhood and lender attributes is to employ a fixed effects model.

To control for time-invariant neighborhood heterogeneities, the tract level analysis estimates models with tract fixed effects. Analogously, the lender level analysis includes both tract and lender fixed effects in order to isolate the impact of unobserved neighborhood and lender heterogeneities that are correlated with neighborhood and lender characteristics included in the estimation. The tract and lender fixed effects models are estimated using both OLS and Poisson regression.

For the OLS regression, the neighborhood fixed effects models are implemented through mean differencing and the lender fixed effects are controlled for by including a lender-specific dummy variable (D_l) representing each lender (less one). The tract fixed effects model is identified by the changes in neighborhood characteristics between years.[35] The lender fixed effects model is identified by the changes in lender characteristics between tracts and between years.

For the Poisson regression, tract fixed effects are estimated by maximizing conditional likelihood function and lender fixed effects are estimated by including dummy variables. Cameron and Trivedi (1998) provide theoretical foundation for fixed effects models with Poisson distribution. In the tract level analysis the dependent variable NOB_{ty} is assumed to have a Poisson distribution with parameter λ_{ty}.

$$\ln\lambda_{\mathrm{ty}} = \delta_t + \beta X_{ty}$$

Here, δ_t represents tract fixed effects. According to Cameron and Trivedi, these fixed effects can be estimated by including a dummy variable for each tract (less one) in the conventional Poisson regression by maximum likelihood. An alternative method is to maximize the conditional maximum likelihood function, conditioning on the count total $\sum_t NOB_{ty}$ for each tract. For the Poisson model, this yields to conditional likelihood function that is proportional to:

$$\prod_t \cdot \prod_y \cdot \left[\frac{\exp(\beta X_{ty})}{\sum_t \exp(\beta X_{ty})} \right]^{NOB_{ty}}$$

Note that the conditioning eliminates the fixed effects (δt). Furthermore, Cameron and Trivedi (1998) show that the maximization of the unconditional likelihood function and the conditional likelihood function produce identical estimates and the covariance matrix. Therefore, choice of estimation method may be dictated by computational convenience.

The interpretation of the regression coefficient of Poisson estimation differs from the interpretation of the OLS estimation. In the OLS estimation, β_j implies the marginal effect of the *jth* regressor on the conditional mean. For the Poisson estimation, the conditional mean function is

$$E\lfloor \lambda_{ty} \mid X \rfloor = \exp(\delta_t + \beta X)$$

Therefore,

$$\frac{\partial E\lfloor \lambda_{ty} \mid X \rfloor}{\partial X_j} = \beta_j \exp(x_i ' \beta)$$

To calculate the marginal effect, we need the estimated β_j and $\exp(x_i ' \beta)$ However, the value of $\exp(x_i ' \beta)$ is different for different tracts in the tract level analysis[36] and makes the interpretation difficult. A direct interpretation of the estimated β_φ can be obtained by computing the Incident Rate Ratio (IRR), which is simply the exponential of the raw Poisson estimate. The interpretation of IRR is the relative change in the incident rate brought by one-unit change in the independent variable.[37]

3.4. Use of census and HMDA based neighborhood characteristics

To implement tract fixed effects in the tract and lender level analysis, I use mean differencing in the OLS estimation and maximize conditional likelihood function in the Poisson estimation. However, for the OLS or Poisson estimation to work, the tract characteristics included in the sample being used must vary between tracts.[38] Since the variables representing neighborhood characteristics are constructed from 1990 census data, they remain unchanged for any given tract across all years in the sample period.[39] For this reason, we cannot employ a tract fixed effects model using census data. The paper utilizes the Home Mortgage Disclosure Act (HMDA) data to construct neighborhood characteristics variables that exhibit variation between years. The construction of HMDA based variables is described in the data and variable construction section. The HMDA based neighborhood characteristics are used to control for changes in neighborhood attributes over time in the tract and lender level analyses.

To implement lender fixed effects in the lender level analysis, I include lender dummies both in the OLS estimation and in the unconditional maximum likelihood function of the Poisson estimation. However, the identification of both OLS and Poisson requires that each lender level variable to vary between years and/or between tracts. For the four lender characteristics variables (Head Distance, Deposit, Equity and Past CRA Rating), this is easily achieved. For any given lender, each of these characteristics varies between tracts and between years.[40] None of the nine neighborhood characteristics variables in the census data varies between years for any given tract and, therefore, would impose an insurmountable hurdle in implementing a tract fixed effects model. However, this would pose no difficulty implementing a lender fixed effects model. This is because for each lender every neighborhood characteristics vary between neighborhoods within a given year. Therefore, the average characteristics of the neighborhoods in which a particular lender operates varies from any particular neighborhood characteristics. To implement lender fixed effects models in the lender level analysis, I use both census and HMDA based neighborhood characteristics. This provides a way to compare the effects of alternative data sets on location choice. The basic structure of estimation strategy is shown in table 1.

Table 1

Regression Type	Levels of Analysis	Fixed Effects	Implementation Strategy	Dataset Used
OLS	Tract level analysis	Tract fixed effects	Mean differencing	HMDA based variables
	Lender level analysis	Lender fixed effects	Inclusion of dummy variable for each lender less one	Both Census and HMDA based variables
		Tract fixed effects	Mean differencing	HMDA based variables
		Both lender and tract fixed effects	Inclusion of lender dummy and tract mean differencing	HMDA based variables
Poisson	Tract level analysis	Tract fixed effects	Maximizing conditional likelihood function	HMDA based variables
	Lender level analysis	Lender fixed effects	Maximizing unconditional likelihood function with dummy variable for each lender less one	Both Census and HMDA based variables
		Tract fixed effects	Maximizing conditional likelihood function.	HMDA based variables
		Both lender and tract fixed effects	Maximizing conditional likelihood function conditional on count total $\sum_t NOB_{ty}$ with lender dummy.	HMDA based variables

4. Data and variable construction

Five different datasets are combined for the empirical analysis to create the 8-year sample that spans from 1992 through 1999. The datasets are (*1*) 1990 and 2000 census data, (*2*) Home Mortgage Disclosure Act (HMDA)[41] data, (*3*) Branch location information of all depository lenders in Connecticut (*4*) FDIC summary of deposit data and (*5*) Community Reinvestment Act (CRA) ratings data. The summary files of the 1990 and 2000 United States census of Population and Housing data provide information on neighborhood characteristics and tract boundaries in Connecticut.[42] Pursuant to HMDA, the Federal Financial Institution Examination Council (FFIEC) collects Loan

Application Registers (LARs) of regulated lenders. The LAR, reported by each lender's main office contains acceptance or denial information for every mortgage application received by the lender. In addition to denial information, HMDA data also contain basic information about borrower, property, and the neighborhood where the property is located. FDIC summary of deposit data provide two key variables representing the lender characteristics: Deposit and Equity information. CRA rating information was collected from the publicly available FFIEC database.

Branch location data was purchased from a private data collection company named Sheshunoff Inc. It contains detailed branch addresses of every depository lender in Connecticut from 1992 to 1999. The branch location information is geocoded to latitude and longitude. In addition, tract boundaries are used to calculate latitude and longitude of geographic center or centroid of each tract.[43] The branch and centroid locations are used to calculate the dependent variable that counts number of branches within one mile distance of tract centers.[44] The location of branches, lenders' head offices and tract centroids are used to create Head Distance variable. This variable is a lender-tract-year variable that calculates the distance between the center of tract *t* and head office location of the lender *l* that has a branch within a one mile radius of the tract center in year *y*. The descriptive statistics of the variables used in the paper are shown in table 2.

Table 2

Tract Level Analysis		
Year	Number of tracts	
	Census	HMDA
1992	823	793
1993	823	789
1994	823	795
1995	823	794
1996	823	791
1997	823	783
1998	823	788
1999	823	789
Total observations	6584	6322

The tract level analysis uses a tract-year sample and attempts to understand the aggregate location choice of all lenders in a particular neighborhood. The dependent variable for this analysis is the number of branches of all lenders that are within 1-mile radius of tract center. This is not the same as the total number of branches of all lenders that are located within the geographic boundary of the tract, or per capital number of branches as used by previous studies (see Caskey 1994 and Avery 1991). The dependent variable used in those studies captures local presence relatively well for smaller or densely populated tracts. However, for larger or thinly populated tracts, total number of branches within tract may not perform well as a proxy for local presence. This is because lenders' service areas rarely coincide with the tract boundaries.[45] Consequently, a lender may choose to locate around the edge of a tract providing its services primarily to surrounding tracts. Antonakes (2001) addresses this issue by using a dependent variable that counts the number of lenders that include a particular tract within their service area. The dependent variable used in this paper addresses this issue by narrowing the area of focus, essentially assuming that if a branch is located within one mile radius of a tract center it necessarily serves the neighborhood. This definition will not count branches that are outside the one-mile perimeter. For smaller tracts, however, this definition may count branches that are located outside the tract but within one-mile radius of the tract center.

Since census based variables representing tract characteristics exhibit no variation between years within one decade, the tract fixed effects model cannot be employed using the census sample. In the paper, the neighborhood characteristics are constructed using information available in HMDA data. Unlike census data, HMDA data is collected annually. Therefore, HMDA based variables remain the same within a given year but vary across different years, allowing us to implement fixed effects models using tract mean differencing for the OLS and maximize conditional likelihood function for the Poisson regression. To construct the population density variable, I use both 1990 and 2000 census information and assume a steady population growth to achieve between year variations.[46]

HMDA based variables are constructed to represent neighborhood characteristics using HMDA application samples for new home mortgage or mortgage refinance submitted to depository and non-depository lenders who are required to report under HMDA. These variables include tract median income, composition of racial groups,[47] percentage male, and percentage of owner occupied housing units. The variables representing percentage old and percentage occupied could not be constructed because information required to create these variables is not available in HMDA data.

For several reasons, use of HMDA data for this purpose is more suitable than any other existing database. First, HMDA data covers almost all depository lenders whose location choice is modeled in this paper. Second, HMDA regulations require that these lenders collect and report key information on a mortgage loan application and the applicant on a loan-by-loan basis. Since HMDA data provides information about applicants who are either residents or potential residents of the neighborhood, it is likely that average characteristics of HMDA applicants would reflect the average characteristics of the residents in the neighborhood. In fact, the key assumption required to construct HMDA based variables and use them to represent neighborhood characteristics is that no systematic difference exists between mortgage applicant pool and the pool of actual residents. For example, if a neighborhood has a large percentage of minority mortgage applicants in a given year, it is reasonable to assume that it would be reflected in the percentage of minority population in that neighborhood.[48]

The samples are organized for tract and lender levels of analysis. Each analysis is conducted using both census and HMDA based variables representing neighborhood characteristics. Table 2 shows the number of tracts in the 8-year sample period. In the dataset with census based variables, the final sample includes identical 823 tracts for all years. Therefore, the sample used for tract level analysis with census based variables has 6584 observations. The tract level analysis using HMDA based variables excludes certain tracts in certain years because either lending activities in those tracts are not HMDA reportable or there was zero mortgage origination in those tracts. The sample for this analysis has 6233 observations as shown in the table above.

Similarly, for lender level analyses, final samples are organized using census and HMDA based variables as shown in table 3. These samples are constructed by combining total number of tracts in each year with every active depository lender in that year for which all relevant lender specific information was available. In both census and HMDA based sample, the number of lenders across different years remains the same. However, the total number of observations differs because fewer tracts available in the HMDA based sample produced smaller numbers of tract-lender combination as shown in the table above.

Table 3

Lender Level Analysis						
Year	Census based variables			HMDA based variables		
	Number of Tracts	Number of Lenders	Number of tract-lender combinations	Number of Tracts	Number of Lenders	Number of tract-lender combinations
1992	823	67	55,141	793	67	53131
1993	823	80	65,840	789	80	63120
1994	823	89	73,247	795	89	70755
1995	823	86	70,778	794	86	68284
1996	823	71	58,433	791	71	56161
1997	823	61	50,203	783	61	47763
1998	823	59	48,557	788	59	46492
1999	823	54	44,442	789	54	42606
	Total observations		466,641	Total observations		448,312

5. Results

The results are presented in two stages. Stage I presents the neighborhood or tract level analysis and stage II presents the lender level analysis. Stage I results are discussed under three headings: (a) Tract level analysis using census data (b) Tract level analysis using HMDA data (c) Tract level analysis with tract fixed effects using HMDA data.

5.1. Stage I: The tract level analysis

5.1.1. Tract level analysis using census data

The tract level analysis using census data is presented in this section. The OLS and Poisson estimates from this analysis are shown in column one and two of table 6 respectively. This analysis estimates the impacts of nine neighborhood characteristics available in census data on the number of branches of all neighborhood lenders that are within a one-mile radius of a tract center (NOB_{ty}). The results find percentage owner occupied (-), population density (+), percentage old (+) and percentage Asian (+) are statistically significant. The neighborhoods with higher percentages of owner occupied housing units are associated with lower branch presence. This impact is significant at 1% significance level and perhaps suggests a revealed preference of the neighborhood residents to keep businesses including bank branches away from their residential locations. The population density also affects branch representations at 1% significance level suggesting greater branch representation in tracts with higher population densities. Branch presence across neighborhoods by all lenders appears to rise with percentage older population at 1% significance level. This is consistent with Chang *et al.* (1997), who postulate that older populations living on fixed incomes or with higher proportion of savings tend to supply larger deposits and thereby attracting financial institutions. Although the result does not find a significant influence of median income on branch presence, the positive relation is consistent with the results of the previous studies.

5.1.2. Tract level analysis using HMDA data

The OLS and Poisson estimates for tract level analysis using explanatory variables constructed from HMDA data is shown in columns three and four of table 6 respectively. As noted in the data and variable construction section, two neighborhood characteristics variables –percentage old, and percentage occupied– could not be created due to non-availability of comparable information in HMDA dataset. However, they are controlled for using census data information to make the analysis comparable to tract level analysis using census data (shown in columns one and two of table 4). Comparing columns three and four with the corresponding

Table 4. Tract Level Analysis. Dependent Variable: NOB_{TY} is the number of branches of all lenders within one-mile radius from the center of tract T in year Y.

Parameter NOB_{TY}	Without Tract Fixed Effect				With Tract Fixed Effect	
	Using Census Data		Using HMDA Data[2]		Using HMDA Data[3]	
	OLS Col 1	Poisson Col 2	OLS Col 3	Poisson Col 4	OLS Col 5	Poisson Col 6
Neighborhood Characteristics						
Tract Median Income	.0162 .0154	.0019 .0037	.0048 .0057	.001 .0013	.003 ** .001	-.0001 .0003
Percentage Black[4]	-.0200 .0141	-.0025 .0020	-.0203 * .0112	-.0005 .0016	-.0106 *** .0025	-.0009 ** .0005
Percentage Asian	.3023 * .1709	.0470 ** .0199	.0331 .0223	.0037 .0027	.0153 *** .0048	.0009 .001
Percentage Other[5]	.0755 .0541	.0027 .0044	-.0080 .0152	.0012 .0019	-.0148 *** .0025	-.0015 *** .0005
Percentage Male	-.0025 .1157	-.0106 .0102	.0221 .0159	.0019 .0018	-.0130 *** .0027	-.0013 ** .0005
Percentage Old	.0895 ** .0408	.0185 *** .0057	.0604 * .0311	.0233 *** .0056	-------	--------
Percentage Occupied	-.0537 .0423	-.0064 .0074	-.0999 ** .0491	-.0159 * .0097	-------	---------
Percentage Owner Occupied	-.0976 *** .0183	-.0227 *** .0025	-.0140 *** .0034	-.0071 *** .0011	-.0048 *** .0010	-.0009 ** .0004
Population Density	.0004 *** .00008	.00003 *** .000007	.0007 *** .00007	.00007 *** .0000086	.0002 *** .00004	.00002 ** .000007

Note: All standard errors are robust to correct for heteroskedasticity and for clustered sampling. Year dummies are not shown.

columns one and two suggests a statistical similarity between census and HMDA datasets in representing neighborhood characteristics relevant in this analysis. None of the explanatory variables changes sign. On the basis of significance, percentage owner occupied and population density continue to be significant at 1% significance level.

5.1.3. Tract level analysis with tract fixed effects using HMDA data

The neighborhood characteristics reported in census data do not vary between years over the sample period.[49] However, the neighborhood characteristics constructed from HMDA data varies between years. Utilizing this variation, this section performs the tract level analysis with tract fixed effects using HMDA based variables. The OLS and Poisson estimates are shown in column five and six of table 4 respectively. By controlling for time-invariant unobserved neighborhood heterogeneities, the estimates of this analysis improve upon the estimates obtained from models without tract fixed effects. However, two neighborhood characteristics[50] vanish from the estimation since they have no variation between years.

Controlling for neighborhood fixed effects causes noticeable changes to the results. This is observed by comparing column five and six with column three and four in table 4 respectively. Neighborhood median income, percentage Black, percentage Asian, percentage other and percentage male[51] are now significant factors in explaining distribution of depository branches across neighborhoods. Percentage owner occupied and population density continue to be significant determinants for branch representation across neighborhoods.

The results suggest that neighborhood fixed effects estimation has absorbed a lot of variation and led to smaller coefficient estimates in general for almost all independent variables included in the model except for percentage male. However, because of reduced variation the fixed effects estimation has also produced more precise estimates. In this case, coefficient estimate and its precision contribute to its significance in opposite ways.[52] In the net, neighborhood income and racial composition become significant factors in explaining branch representations after controlling for time invariant, unobserved neighborhood fixed effects. Some of the important neighborhood unobserved fixed effects are likely to include residential zoning related heterogeneities, unobserved attributes associated with central business districts and unique characteristics of neighborhoods that border with New York state providing tax benefits to those who work in New York city and reside in these neighborhoods.

The neighborhood fixed effects results finds that when the number of owner occupied housing unit doubles, the number of branches in the neighborhood falls by one half of one branch according to the OLS estimates. According to the Poisson estimates, a 1% increase in percentage owner occupancy in a neighborhood reduces the existing number of branches in the neighborhood by a multiple of .9991,[53] which is a .09% decline.

The standard approach in the literature is to employ a tract level analysis to analyze branch representations across neighborhoods. Consistent with this approach, this paper finds that neighborhood characteristics such as median income and racial composition affect the branch representation of financial institutions as found in several previous studies.[54] Although these analyses point out an important public policy concern, where a faction of poor and minority population may have limited access to basic depository services based on where they reside and may be forced to rely on fringe banking, these analyses are incomplete. While pointing out the tracts that lack the access to financial services, these analyses do not examine if this outcome is a result of depository lenders decisions to stay away from low-income and minority neighborhoods. Lender level analysis presented in stage II attempts to understand the factors that affect these lenders' branch location choices.

5.2. Stage II. The lender level analysis

The lender level analysis estimates models of branch location choice of individual lenders using both neighborhood and lender characteristics. This analysis is presented in stage II and discussed under four headings: (*a*) Lender level analysis with neighborhood characteristics using census data, (*b*) Lender level analysis with neighborhood characteristics using census data and additional lender level controls, (*c*) Lender level analysis with neighborhood characteristics using census data and additional lender level controls, and lender fixed effects, (*d*) Lender level analysis with both lender and tract fixed effects using HMDA data.

5.2.1. Lender level analysis with neighborhood characteristics using census data

This analysis uses the same nine neighborhood characteristics from the census data that was used in the tract level analysis in the previous section. Now we would like to know how these variables affect individual lender's branch location choice. The dependent variable for this analysis is the number of branches of individual lenders that are within a one mile distance of a neighborhood center.[55] The results of this analysis are presented in column three and four of table 5. Results of the tract-level analysis using census data from table 6 are reproduced in column one and two of table 5 in order to compare the impacts of explanatory variables on location choice of individual lenders (lender-level analysis) and branch representation of all lenders across neighborhoods (tract-level analysis). Therefore, the difference between the first two columns with the corresponding second two columns is attributable to the change of dependent variable from tract specific NOB_{ty} to lender-tract specific NOB_{lty}.

Table 5. Comparison of Tract Level and Lender Level Analysis using the same Neighborhood Characteristics as Explanatory Variables. Dependent Variable: NOB_{TY} is the number of branches of all lenders within one-mile radius from the center of tract T in year Y. Dependent Variable: NOB_{LTY} is the number of branches of individual lenders within one-mile radius from the center of tract T in year Y.

Parameter	Using Census based Neighborhood Characteristics			
	Tract Level Analysis (Reproduced from col.1 and col.2 of table 3)		Lender Level Analysis	
	Dependent Variable: NOB_{TY}		Dependent Variable: NOB_{LTY}	
	OLS Col 1	Poisson Col 2	OLS Col 3	Poisson Col 4
Neighborhood Characteristics				
Tract Median Income	.0162 .0154	.0019 .0037	.00008 .0001	.0002 .0036
Percentage Black	-.0200 .0141	-.0025 .0020	-.0003 *** .0001	-.0062 *** .0022
Percentage Asian	.3023 * .1709	.0470 ** .0199	.0020 .0012	.0405 ** .0192
Percentage Other	.0755 .0541	.0027 .0044	.0002 .0003	-.0015 .0044
Percentage Male	-.0025 .1157	-.0106 .0102	-.0001 .0007	-.0123 .0086
Percentage Old	.0895 ** .0408	.0185 *** .0057	.0006 ** .0002	.0184 *** .0056
Percentage Occupied	-.0537 .0423	-.0064 .0074	-.0003 .0003	-.0043 .0072
Percentage Owner Occupied	-.0976 *** .0183	-.0227 *** .0025	-.0007 *** .0001	-.0236 *** .0023
Population Density	.0004 *** .00008	.00003 *** .000007	.000002 *** .000006	.00003 *** .000007

Table 6. Lender Level Analysis with Additional Lender Controls. Dependent Variable: NOB_{LTY} is the number of branches of individual lenders within one-mile radius from the center of tract T in year Y.

Parameter NOB_{LTY}	Using Census based Neighborhood Characteristics			
	Lender Level Analysis (Reproduced from Col 3 and 4 of table 4)		Lender Level Analysis with Additional Controls: Without Lender and Without Tract Fixed Effects	
	OLS Col 1	Poisson Col 2	OLS Col 3	Poisson Col 4
Neighborhood Characteristics				
Tract Median Income	.00008 .0001	.0002 .0036	.0002 ** .0001	.0022 .0039
Percentage Black	-.0003 *** .0001	-.0062 *** .0022	-.0003 *** .0001	-.0079 *** .0022
Percentage Asian	.0020 .0012	.0405 ** .0192	.0026 ** .0012	.0443 *** .0169
Percentage Other	.0002 .0003	-.0015 .0044	.0001 .0003	-.0069 .0043
Percentage Male	-.0001 .0007	-.0123 .0086	-.0001 .0006	-.0101 * .0053
Percentage Old	.0006 ** .0002	.0184 *** .0056	.0005 ** .0002	.0141 *** .0052
Percentage Occupied	-.0003 .0003	-.0043 .0072	-.0006 ** .0002	-.0114 * .0058
Percentage Owner Occupied	-.0007 *** .0001	-.0236 *** .0023	-.0008 *** .0001	-.0243 *** .0022
Population Density	.000002 *** .000006	.00003 *** .000007	.000003 *** .0000006	.00002 *** .000007
Lender Characteristics				
Head Distance	---------	---------	-.0014 *** .00007	-.0681 *** .0021
Deposits	---------	---------	.0936 *** .0058	.7330 *** .0310
Equity	---------	---------	-.1540 *** .0293	-1.850 *** .3090
Past CRA Rating	---------	---------	-.0029 ** .0015	-.1089 * .0385

The results suggest that lender level analysis is comparable to tract level analysis. The signs of the coefficients for all independent variables remain unchanged implying that these neighborhood attributes affect individual lenders´ location choice in a similar fashion as they affect aggregate location choice of all lenders. However, the magnitude of their impacts declines because the coefficients in lender level analysis measure the marginal effect of each covariate on the conditional mean number of branches of each lender as opposed to the total number of branches of all lenders.

5.2.2. Lender level analysis with neighborhood characteristics using census data and additional lender level controls

The lender level analysis with tract level controls is extended by including four lender characteristics into the set of explanatory variables. The additional variables are distance of lender's head office from neighborhood center (Head Distance), deposits available to the lender (Deposit), equity held by the lender (Equity) and average of three past CRA ratings (Past CRA Ratings). The OLS and Poisson estimates from this analysis are presented in column three and four of table 6. To compare, the results with the lender level analysis without additional lender controls are reproduced from column three and four of table 5 to column one and two of table 6. The difference between the first two columns with the corresponding second two columns is attributable to the inclusion of four lender specific controls.

The inclusion of lender characteristics changes the significance of several independent variables that were present in the previous analysis with only nine tract level controls. However, the direction of correlation, magnitude and standard error of all common variables in both analyses remain stable. The neighborhood median income, percentage Asian and percentage occupied become significant factors in explaining branch location choice of individual depository lenders in OLS regression. Similarly, percentage Asian and percentage occupied become significant in both OLS and Poisson estimations.

Among the lender characteristics, while the head office location, lender's equity and poor CRA ratings reduce branch presence of individual lenders, deposit has an opposite effect. The impact of the head office location, lender's equity and deposits are strongly significant at a 1% significance level. Past CRA ratings is significant at a 5% significance level.[56] This paper finds that branch presence is higher in the neighborhoods that are close to lenders' head office. The divergent impacts of deposit and equity suggest that while deposits increase branch presence, equity might have the opposite effect. Given a substantial degree of risk and uncertainty associated with branch openings,[57] it is possible that managers in lending institutions are willing to assume higher risk when they have more deposits, which essentially are debts. Conversely, they might be willing to accept lower risk when banking assets are in the form of equity. This finding indicates a possibility of divergent impacts of capital structure on risk tolerance of the managers at the depository institutions with regard to branch openings. Higher past CRA rating is associated with lower branch representation suggesting that lenders with poor CRA performance have experienced reduced branch representation.[58] This reduced branch presence might explain one of the findings of Dahl, Evanoff and Spivey (2000) that public and regulatory pressure exerted through poor CRA ratings did not lead to higher low-income mortgage lending. Although the significance of past CRA ratings on the number of branches disappears, the sign remains negative in the subsequent analysis with lender and tract fixed effects.

5.2.3. Lender level analysis with neighborhood characteristics using census data and additional lender level controls and lender fixed effects

In this section, a lender fixed effects model is estimated using nine neighborhood characteristics from census data and four lender characteristics. The results are shown in table 7. To distinguish the impact of lender fixed effects, the lender level analysis without the lender fixed effects is reproduced from column three and four of table 8 to column one and two of table 7.

The results indicate that the inclusion of lender fixed effects has almost no impact on the estimates of neighborhood characteristics variables. However, controlling for time-invariant, unobserved lender heterogeneities produces noticeable change in the lender level variables. First, the coefficient estimate for lender's equity rises and becomes positive in the OLS regression maintaining a 1% level of significance. The depository lenders differ from each other with regard to the regulatory agency that supervises them and imposes capital adequacy requirements

Table 7. Lender Level Analysis with Additional Controls and Lender Fixed Effects using Census Data. Dependent Variable: NOB_{LTY} is the number of branches of individual lenders within one-mile radius from the center of tract T in year Y.

Parameter NOB_{LTY}	Using Census based Neighborhood Characteristics			
	Without Lender and Without Tract Fixed Effects (Reproduced from Col 3 and 4 of table 5)		With Lender Fixed Effects and Without Tract Fixed Effects	
	OLS Col 1	Poisson Col 2	OLS Col 3	Poisson Col 4
Neighborhood Characteristics				
Tract Median Income	.0002 ** .0001	.0022 .0039	.0002 ** .0001	.0014 .0037
Percentage Black	-.0003 *** .0001	-.0079 *** .0022	-.0003 *** .0001	-.0089 *** .0023
Percentage Asian	.0026 ** .0012	.0443 *** .0169	.0027 ** .0012	.0384 ** .0166
Percentage Other	.0001 .0003	-.0069 .0043	.0001 .0003	-.0069 .0053
Percentage Male	-.0001 .0006	-.0101 * .0053	-.0001 .0006	-.0099 * .0053
Percentage Old	..0005 ** .0002	.0141 *** .0052	..0005 ** .0002	.0126 ** .0052
Percentage Occupied	-.0006 ** .0002	-.0114 * .0058	-.0006 ** .0002	-.0110 * .0058
Percentage Owner Occupied	-.0008 *** .0001	-.0243 *** .0022	-.0008 *** .0001	-.0243 *** .0022
Population Density	.000003 *** .0000006	.00002 *** .000007	.000003 *** .0000006	.00002 *** .000007
Lender Characteristics				
Head Distance	-.0014 *** .00007	-.0681 *** .0021	-.0016 *** .00008	-.0618 *** .0022
Deposits	.0936 *** .0058	.7330 *** .0310	.0082 *** .0017	.1440 *** .0289
Equity	-.1540 *** .0293	-1.850 *** .3090	.1170 *** .0268	-.7470 *** .0000002
Past CRA Rating	-.0029 ** .0015	-.1089 * .0385	-.0001 .0008	-.0374 .0252

upon them. Higher capital requirements increase observed equity available to lenders but decrease cash flows required to expand branch networks. Not controlling for unobserved lender heterogeneities such as regulatory controls on capital requirements that remain unchanged over time are likely to negatively bias the impact of lender's equity variable in models without lender fixed effects. The lender fixed effects model finds that similar to deposits, equity also increase branch presence according to the OLS estimation. Second, the impact of past CRA ratings falls and becomes insignificant in both OLS and Poisson regression. This suggests that higher (poor) CRA ratings are also correlated with time-invariant, unobserved lender heterogeneities, perhaps including these lenders´ CRA related efforts and commitments[59] that remain in place for multiple years. When this and other unobserved lender heterogeneities are controlled for the impact of the CRA ratings disappears suggesting that poor CRA ratings have no significant impact on branch location choice of individual lenders.

5.2.4. Lender level analysis with both lender and tract fixed effects using HMDA data

Finally this section includes tract fixed effects in addition to lender fixed effects to control for time-invariant unobserved neighborhood heterogeneities. However, this cannot be accomplished using census based neighborhood variables because tract attributes in census data have no variations between years for a given tract. The neighborhood characteristics constructed from HMDA disclosure information exhibit this variation and these characteristics are used to implement a neighborhood fixed effects model. Columns five and six of table 8 present results of lender level analysis that include both lender and tract fixed effects using HMDA based variables. To distinguish the impact of lender and tract fixed effects, analysis with neither lender nor tract fixed effects is presented in columns one and two of table 8 and analysis with lender fixed effects but without tract fixed effects is shown in columns three and four of table 8. For comparison purpose, all variables in this table are HMDA based except for percentage old and percentage occupied, which are based on census data.[60] However, these two variables vanish from the tract fixed effects estimation. The results presented in table 8 also corrects for robust, clustered standard errors.

The inclusion of the tract fixed effects in the lender level analysis produces striking results. Among the neighborhood characteristics, the

Table 8. Lender Level Analysis with Lender and Tract Fixed Effects using HMDA data. Dependent Variable: NOB_{LTY} is the number of branches of individual lenders within one-mile radius from the center of tract T in year Y

Parameter NOB_{LTY}	Using HMDA based Neighborhood Characteristics					
	Without Lender and Without Tract Fixed Effects		With Lender Fixed Effects and Without Tract fixed effects		With Lender and with Tract Fixed Effects	
	OLS Col 1	Poisson Col 2	OLS Col 3	Poisson Col 4	OLS Col 5	Poisson Col 6
Neighborhood Characteristics						
Tract Median Income	.0001 ** .00004	.0010 .0014	.0001 ** .00004	.0004 .0014	-.000008 .00003	-.0001 .0006
Percentage Black	-.0003 *** .00008	-.0036 ** .0018	-.0002 *** .00008	-.0042 ** .0017	-.00003 .00004	.00003 .0007
Percentage Asian	.0004 ** .0002	.0065 ** .0033	.0004 ** .0002	.0050 * .0030	.0001 .00009	.0030 ** .0015
Percentage Other	-.00008 .0001	-.0007 .0021	-.00008 .0001	-.0015 .0020	.00001 .00004	.0008 .0008
Percentage Male	.000001 .0001	-.0005 .0021	.000001 .0001	-.00007 .0021	-.0001 *** .00005	-.0019 ** .0008
Percentage Old	---------	---------	---------	---------	---------	---------
Percentage Occupied	---------	---------	---------	---------	---------	---------
Percentage Owner Occupied	-.0001 *** .00002	-.0071 *** .0011	-.0001 *** .00002	-.0076 *** .0011	-.00006 *** .00002	-.0035 *** .0006
Population Density	.000005 *** .0000005	.00006 *** .000009	.000005 *** .0000005	.00006 *** .000008	-.0000002 .0000008	-.00006 *** .00001
Lender Characteristics						
Head Distance	-.0014 *** .00007	-.0663 *** .0023	-.0016 *** .00008	-.0618 *** .0024	-.0018 *** .00002	-.0654 *** .0006
Deposits	.08330 *** .0005	.5550 *** .0285	.0074 *** .0019	.1560 *** .0310	.0074 *** .0025	.5470 *** .0313
Equity	-.1070 *** .02.85	-.0021 .2880	.1150 *** .0277	-.9060 *** .2240	.1150 *** .0175	-1.950 *** .2340
Past CRA Rating	-.0048 *** .0016	-.2704 *** .0408	.0002 .0009	-.0424 .0265	.0001 .0015	-.3083 *** .0329

Notes: 1. Home Mortgage Disclosure Act (HMDA) based variables are constructed to represent neighborhood characteristics using HMDA applications for new home mortgage or mortgage refinance submitted to depository and non-depository lenders who are required to report under HMDA. 2. Percentage Old and Percentage Occupied are two tract level variables that could not be created using information reported in the HMDA sample. 3. Tract fixed effect cannot be implemented using the Census data. 4. Percentage White is the omitted category. 5. Percentage Other includes Hispanic and American Indian population.

evidence of significant impact of median income or racial composition[61] on the branch location choice of individual lenders disappears. Because of reduced variation associated with fixed effect estimation both magnitude and standard error of the coefficient estimate fall for median income, percentage Black and percentage Asian. After controlling for unobserved, time-invariant neighborhood fixed effects, the magnitude of coefficient falls much more than the precision resulting in no evidence of significant impact of these variables on branch location choice at the lender level.

A comparison of the lender level analysis shown in table 8 with the tract level analysis shown in table 4 reveals an important fallacy of composition in which what is true for an individual lender is not true for the group in which the lender belongs. Specifically, the tract level analysis with neighborhood fixed effects[62] finds that neighborhood income and racial composition are significant determinants of branch representation. In the lender level analysis, however, there is no evidence that these two variables affect branch location choice of individual lenders. This suggests an incongruity between neighborhood and lender level analysis, where neighborhood median income and racial composition do not affect individual lenders' branch location choice, but they affect the distribution of branches across tracts by all lenders in aggregate.

This paper proposes that this incongruity is related to unobserved neighborhood heterogeneities that fulfill two properties simultaneously. First, these neighborhood heterogeneities are correlated with neighborhood income and racial composition. Second, they affect individual lenders' branch location choice. This incongruity can be explained by using the concept of disparate-treatment and disparate-impact in the context of branch location choice. Based on income or race, when an individual lender adversely treats a neighborhood by staying away from it, we may think of this behavior as an incidence of disparate treatment. On the other hand, whenever certain neighborhoods observe adverse distribution of branches even though no disparate treatment is practiced, we may regard this as a disparate-impact or disparate-outcome of branch location choices. Therefore, another way to express this incongruity is to ask how an absence of disparate treatment toward certain neighborhoods can still produce disparate impact in these neighborhoods.

This paper suggests that this apparent incongruity happens because the omitted neighborhood attributes that affect individual lenders' location choice are also correlated with neighborhood income or racial composition. For example, in the lender level analysis we find that lower deposits are associated with fewer branch representations. When a lender makes branch location choice based on availability of deposits, no direct disparate-treatment against low-income or minority dominated neighborhood is exercised. However, a lower deposit is likely to be associated with the neighborhoods with low median income and high minority population. Therefore, despite no disparate-treatment low-income and minority neighborhoods will experience fewer branch representations –a disparate impact. By conducting a lender level analysis in conjunction with the traditional tract level analysis, we are able to distinguish disparate treatment from disparate impact and obtain a more complete understanding of depository lenders' branch location choice and aggregate outcome of these choices on branch distribution across neighborhoods. The paper points out poor and minority neighborhoods may continue to observe inadequate representations of depository branches and lack of basic financial services even when lenders do not stay away from these neighborhoods. The tract level analysis in previous studies fails to detect this.

Among other neighborhood characteristics variables, a robust impact is found for the percentage owner occupancy, which remains negative and significant at a 1% level of significance under both OLS and Poisson specifications and across both census and HMDA based variables in the lender level analysis.[63] As the number percentage of owner occupied housing units in a neighborhood increases, the neighborhood becomes more residential and its residents gain control over neighborhood's zoning and land use decisions. It is likely that these residents would drive businesses including lending institutions away from their neighborhood. This location specific attribute or greater control of residents over land uses is correlated with higher level of percentage owner occupancy. This unobserved attribute also affects branch presence in the neighborhood and is hence correlated with the error term in the estimating equation. Not controlling for this neighborhood characteristics would bias the estimates of the percentage owner occupancy variable in models without tract fixed effects. After controlling for omitted neighborhood fixed effects, the parameter estimate changes from -.0001 to -.00006 in the OLS estimation. This result implies that doubling the percentage owner occupancy in a neighborhood reduces the number of branches by .006 of one branch for any given lender. According to the Poisson estimates, a one percent increase in neighborhood percentage owner occupancy reduces the existing number of depository branches by a multiple of .9965, which is close to a .35% decline.

Among the variables representing lender characteristics, the head office distance and deposits show a robust impact across all specifications. The number of branches rises in the neighborhoods for the lending institutions that are closer to lenders' head office location. The results from the lender level analysis suggest that a 100 mile increase in the neighborhood to head office distance reduces the number of branches for the lender by $1/5^{th}$ of one branch according to OLS estimation. According to the Poisson estimate, a one mile increase in head office location reduces the existing number of branches by a multiple of .936, which is a 6.33% decline of existing branches. The number of branches rises with available deposits for individual lenders across all specifications. The lender level analysis shows that a lender opens up a new branch as its available deposits increase by 135 million according to the OLS estimate.

6. Conclusion

This paper examines the factors that affect the distribution of branches of depository lenders using tract level analysis with neighborhood fixed effects and branch location choices of individual lenders using lender level analysis with both lender and neighborhood fixed effects. The paper utilizes the Home Mortgage Disclosure Act (HMDA) information to construct neighborhood characteristics that vary across years. Exploiting this across-time variation in HMDA-based neighborhood characteristics, neighborhood fixed effects models are estimated. After controlling for unobserved neighborhood heterogeneity using tract fixed effects, the tract level analysis finds that median income, proportion of minority population[64] and population density significantly influence branch representation. However, in the lender level analysis, the paper does not find evidence of the impact of income or race on branch location choice of individual lenders once neighborhood fixed effects are included.

The paper proposes that this apparent incongruity between lender and neighborhood level analysis can be compared to disparate treatment toward and disparate impact on low income and minority dominated neighborhoods with regard to branch location choice. Specifically, although the paper finds no evidence of impact of neighborhood income or race on the location choice of individual lenders (no disparate treatment), income and race significantly affect the aggregate distribution of branches across neighborhoods (disparate impact). In the context of this paper, unobserved neighborhood heterogeneities or market conditions that affect individual lender's branch location choice are at the same correlated with neighborhood income and racial composition. In the lender level analysis, this is detected by controlling for time constant neighborhood fixed effects. This cannot be detected in the tract level analysis even after controlling for neighborhood fixed effects because lender behavior is not modeled in the tract level analysis. The lender level analysis shows that neighborhoods may continue to remain underserved by depository institutions even when lenders do not make their location choice based upon neighborhood income or racial composition. This underscores the importance and advantage of the lender level analysis over traditional tract level analysis.

Among neighborhood characteristics, the paper finds a significant and robust negative impact of percentage owner occupied housing units on branch location choice of depository lenders. This variable is significant at a 1% significance level in both tract and lender level analysis under all specifications. The negative impact of percentage owner occupancy is likely to be associated with unobserved zoning and land use regulations related to neighborhood fixed effects that influence branch presence. As percentage of owner occupied housing units in a neighborhood increases, the neighborhood becomes more residential and its residents gain control over the neighborhood's zoning and land use decisions. It is likely that these residents would drive businesses including lending institutions away from their neighborhood. After controlling for these unobserved neighborhood heterogeneities, the paper finds that when the number of owner occupied housing units double the number of branches in the neighborhood falls by one half of a branch in the tract level analysis and .006 of one branch in the lender level analysAmong the lender characteristics, the paper finds significant and robust effects of head office location and deposits available. These two variables affect branch location choice of depository lenders at a 1% significance level under all specifications. While farther head office location reduces branch presence, larger amounts of deposits increases the same. A 100 mile increase in the neighborhood to head office distance reduces the number of branches for the lender by $1/5^{th}$ of one branch. This finding is consistent with Felici and Pagnini [2004], who find pre-existing location of lenders is an important determinant of new branch location. With regard to deposits, the paper finds that lenders open up a new branch when available deposits increase by 135 million dollars.

From a policy perspective, reduced branch presence in neighborhoods that are farther from lenders' head office location may have implications for consolidations in the banking industry. Research finds that consolidations through mergers and acquisitions provide scale efficiency [Kwan, 2004]. However, not all consolidations affect the head office location in a similar fashion. For example, mergers of several local banks may not affect their head office distance in any substantial way. However, acquisition by a large, distant bank often has a substantial impact on head office distance. The consolidations that substantially increase head office distance may remove decision making authority away from local banks because lending decisions are typically made at lenders' head offices. These consolidations can adversely affect loan growth especially for small business loans, where lender-borrower relationships play a crucial role for underwriting. This paper provides an additional reason for lower loan growth following such consolidations. According to this paper whenever consolidations increase head office distance for certain lenders, the neighborhoods served by those lenders are likely to observe lower branch representations. The lower branch representations in turn would contribute to lower loan growth. This conjecture is consistent with the Avery and Samolyk (2003) that found that consolidation activities involving larger banks were associated with lower loan growth, whereas community bank consolidations resulting in a greater presence of community banks were associated with higher loan growth. This suggests that while weighing the costs and benefits of consolidations, out-of-state and distant mergers should be viewed and treated differently.

7. Notas

1 See Immergluck (2004) and Caskey (1994).
2 See Stegman (2003), Barr (2003), and Seidman and Tescher (2004).
3 For a study that uses data from New York State, see Chang *et al*. (1997). Connecticut data used in this paper shows that in the years 1992 through 1999, 26 percent of branches were located in the top 5 percent of tracts with the highest branch concentration. On the contrary, the bottom 10% had only 0.53 percent of all branches.
4 See Kiser (2002), Kutler (1996) and Caskey (1991).
5 See Avery (1991), Caskey (1994) and Antonakes (2001).
6 See Cohen and Mazzeo (2004), Felici and Pagnini (2004).
7 Neither disparate treatment nor disparate impact with regard to branch location choice based on income, race or ethnicity is illegal or prohibited under any federal laws. However, Community Reinvestment Act (CRA) may indirectly address such behaviors of lenders when lenders seek for federal approval on the ground that lenders failed to meet the credit needs of the entire community.
8 Avery (1991) and Caskey (1994). The related papers are Chang *et al*. (1997), Antonakes (2001), Felici and Pagnini (2004), Cohen and Mazzeo (2004), Kiser (2002).

9 Avery study included commercial banks, thrifts institutions, check cashing companies and loan and mortgage companies.
10 Atlanta, Boston, Cleveland, Detroit and Philadelphia.
11 Annual household income of 20,000 and below in 1989 dollar.
12 50 percent or more.
13 Included control variables were median home values, per capita number of owner occupied homes, number of firms and employees, percentage of employed residents with white-collar jobs and dummies for center city and the city.
14 Atlanta, Denver, New York City, San Jose and Washington D.C.
15 Obermiller (1988) and Dymski, Veitch and White (1991).
16 The term rational herding has been used to describe situations in which it is individually rational for agents/firms to imitate the actions of others even though such behavior can lead to sub-optimal aggregate outcome. For illustration, authors utilizes information externality model proposed by Lang and Nakamura [1993] to show one of the ways in which bank branch location choice yields rational herding.
17 The neighborhood controls include total population, fraction of non-white population, population over 65, fraction of high school graduates, fraction of poor households, fraction of rental units, median house value of owner occupied units and number of working population.
18 The dependent variable in the Chang *et al.* [1994] paper is the number of new branch openings in a given year and the key independent variable to test for herding behavior is the total number of pre-existing branches. Since the independent variable is a function of lagged dependent variable, the authors point out that a tract fixed effects model will produce inconsistent estimates. In addition, the paper could not employ a neighborhood fixed effects model because census data used in the paper had no variation between years.
19 Title VIII of the Housing and Community Development Act of 1977.
20 Assessment area is a lender specific service area reflecting lenders' type, size, capabilities, branch locations and geographic distribution of loans and deposits. For a detailed review CRA regulations relating to assessment area see Hossain (2004).
21 Besides banking regulators, Department of Justice (DOJ) has the authority to intervene when a pattern of geographic discrimination in branching

is detected. In a 1994 landmark decision, DOJ issued a consent decree to Chevy Chase Federal Savings Bank in Maryland that had most of its branches in relatively affluent neighborhoods of Washington D.C. According to the consent, the bank agreed to open several branches in minority neighborhoods.

22 As Chang *et al.* cite that lender may commission a private market research company to gather information about profit potentials before branch openings.

23 The multi-market bank, single market bank and thrifts.

24 Similarly, Amel and Hannan (1999) finds that the price differential often does not justify switching of banks.

25 These services typically include basic deposit service, check cashing, payment service, savings account, consumer and commercial loans.

26 These services may include cashier's check, bank draft, money order, notary public, safety deposit, provision bank statements etc.

27 Although similar in purpose, NOB_{ty} is not same as number of branches in a tract or per capita number of branches as used in several previous studies (Caskey 1994 and Avery 1991). Since lenders' service area and tracts' geographic boundary rarely coincide, it is possible for a lender to locate its branch at the edge of a large tract and get counted in the number of branch count. However, this branch may primarily serve other surrounding tracts. To capture the provision of banking services through branch representation, NOB_{ty} is designed with a variable degree of focus by changing radius from tract center. The results presented in this paper are based on the number of branches that are within one mile distance of tract center. Therefore, radius is one mile. This point is further elaborated in Data and Variable Construction section.

28 Percentage White is regarded as the omitted category.

29 Percentage female is regarded as the omitted category.

30 Percentage old includes population of 65 and older.

31 The subscripts *l*, *t* and *y* represent lender, tract and year respectively.

32 This comparison will be possible for both OLS and Poisson regression.

33 See Felici and Pagnini (2004), and Cohen and Mazzeo (2004).

34 See Henderson (2001), Li, Hossain and Ross (2003).

35 Since neighborhood characteristics in the census data have no between tract variation, tract fixed effects model is estimated using neighborhood characteristics constructed from Home Mortgage Disclosure Act (HMAD) data. The construction of HMDA based variables are described in the following section.

36 In the lender level analysis, the value of the expression exp($x_i'\beta$) will be different for different lenders.

37 To see how IRR works, lets compute the predicted *rate of occurrence* (or incidence rate) when one the independent variables (x_i) change from 0 to 1. This rate will be

$$\frac{E[Y_i \mid x_i = 1]}{E[Y_i \mid x_i = 0]} = \frac{\exp[\beta_{i \neq j} X_{i \neq j} + \beta_j (1)]}{\exp[\beta_{i \neq j} X_{i \neq j} + \beta_j (0)]} = \exp(\beta_j)$$

Therefore, the relative change in the incident rate brought by one-unit change in the independent variable is the exponential of the Poisson coefficient. Interpretation of incident rate ratio (IRR) is substantially different from OLS estimate. While OLS estimate can take positive (increasing impact), zero (no impact) and negative (decreasing impact) value, IRR is always non-negative. This is because exponential is a non-negative function. When the IRR of an independent variable is more (less) than 1, the impact of the independent variable on the dependent variable increases (decreases) by a factor of the IRR estimate. When the IRR is equal to one, the independent variable has no impact on the dependent variable.

38 In the absence of tract level variation, tract mean will be exactly same as the value of each observation, producing only zeros in the mean differenced sample.

39 The sample period (1992-1999) of this paper falls within a decade. Therefore, every census tract characteristics is same across all years in the sample except for the population density variable that uses both 1990 and 2000 census data. Construction of population density variable is described in the data and variable construction section.

40 In addition, no lender operates in all 823 census tracts. Therefore, the average characteristics of any given lender across all years and tracts will differ from the characteristics of that lender in any particular tract-year combination. The average lender characteristics will also differ because not

all lenders remain active in all the years in the sample period. Restructuring of lending institutions through mergers, consolidations and bank closings during the sample period makes the variation possible.

41 Home Mortgage Disclosure Act (HMDA): The Congress enacted the Home Mortgage Disclosure Act (HMDA) in 1975. The goal of HMDA was to provide sufficient information to determine whether depository institutions are filling their obligations to serve the housing needs of the communities and neighborhoods in which they are located. (12 USC 2801(b)) HMDA requires depository institutions and their subsidiaries to report information on the total number and dollar value of mortgage originated and purchased, housing attributes, applicant race, income and neighborhood characteristics at each loan application level.

42 Except population density, all other neighborhood characteristics variables were based on 1990 census. Population density variable used both 1990 and 2000 census to create a weighted average population density that by design changes over the years at a constant rate. This across year variation was essential to set up a tract fixed effects model.

43 In GIS terminology, tract center is referred to as tract centroid, which is the geographic center of a polygon. Tract center often is not the center highest population density. Average tract size in Connecticut was about 2.5 miles in the 1990 census.

44 For the tract level analysis the dependent variable is NOB_{ty} that counts the number of branches of all lenders that are located within one mile radius of the centroid of tract t in year y. For the lender level analysis the dependent variable NOB_{lty} that counts the number of branches of lender l that are within one mile radius of the centroid of tract t in year y.

45 Service are often includes multiple contiguous tracts. This creates possibilities in which a lender may provide services without locating in the tract or a lender may locate in a tract without including a large part of the tract within its service area.

46 Since the dataset used in this paper spans a period (1992-1999) that falls between these two census years, I construct a weighted average population density for each year. Underlying assumption to create this variable is that the change in population density from 1990 to 2000 occurs at a constant rate.

47 Percentage Black, percentage Asian and percentage Hispanic keeping percentage White as the omitted category.

48 Since HMDA reports borrower characteristics, this correlation would be higher for neighborhoods with more owner occupied units. Therefore, I conduct the tract fixed effects analysis by dropping tracts with fewer owner occupied (large number of renters) housing units as a robustness check. The robustness check does not alter the results obtained from the full sample.

49 Census data is collected once in every ten years. The sample period for this paper includes 1992 through 1999, which falls between 1990 and 2000 census years.

50 Similar to previous analysis, percentage old and percentage occupied are controlled for using census data. However, they vanish from the estimation in the tract fixed effects model since they have no within tract variation between years.

51 Percentage Asian and Median income is found to be significant only in OLS estimation.

52 While the smaller coefficient estimate contributes to reduced significance, lower precision increases the significance.

53 This is also known as Incident Rate Ratio (IRR) obtained by taking exponential of the Poisson estimate.

54 See Avery (1991) and Caskey (1994).

55 Only difference between this analysis and the tract level analysis using census data is in the dependent variable. In this analysis the dependent variable is number of branches of individual lenders as opposed to number of branches of all lenders within one mile of tract centroid.

56 However, significance of this variable disappears in the subsequent analysis in which lender and tract fixed effects are included.

57 Chang *et al*. (1997) describes this uncertainty in terms of profitability associated with branch openings. In addition to uncertain profitability, branch openings involve substantial direct initial set up cost and indirect cost of closing a branch. Closing a branch not only requires federal approval but also risks negative publicity stemming from the opposition of community groups and activists. This high cost and uncertainty make branch opening a risky investment.

58 Descriptive explanation of the CRA ratings are as follows:
1 = outstanding, 2 = satisfactory, 3 = needs to improve and 4 = substantial noncompliance.

59 For more on CRA agreements, see Bostic and Robinson (2003) and Hossain (2004).

60 The lender level analysis *with neither lender nor tract fixed effects* and *with lender but without tract fixed effects* as shown in first four columns in table 10 can be performed using variables constructed from both census (table 9) and HMDA data (table 10). However, the analysis *with lender and with tract fixed effects* as shown in last two columns of table 10 must be conducted using HMDA based variables. Therefore, for comparison purpose all analyses presented in table 10 use variables constructed from HMDA data. However, a comparison of the first two analyses in table 9 and 10 reveals the effect of using variables constructed from different datasets. This comparison shows census and HMDA based neighborhood characteristics produce statistically similar results and stable estimates.

61 Estimate for Asian in the Poisson estimation continues to remain significant at 5% level of significance.

62 This analysis is presented in the last two columns of table 6.

63 The percentage owner occupied is also significant at 1% significance level in the tract level analysis as observed in table 6.

64 Black, Asian and Hispanic are considered while keeping white as the reference group.

8. References

Amel, D. and T. Hannan (1999). "Establishing banking market definitions through estimation of residual deposit supply equations." *Journal of Banking and Finance* 23, 11, pp. 1667-90.

Amel D. F. and Liang J. N. (1997). "Determinants of entry and profits in local banking markets." *Review of Industrial Organization*, 12, 1, pp. 59-78.

Antonakes, S. L. (2001). "Assessing the Community Reinvestment Act: Impact on low income and high minority communities." *The Journal of Business and Economic Studies*, 7, 1, pp. 1-31.

Avery, R. (1991). "Deregulation and the location of financial institution offices." *Economic Review* (Federal Reserve Bank of Cleveland), 27, 3.

Avery, R., Bostic, R., Calem P. and Canner, G. (1997). "Changes in the distribution of banking offices." *Federal Reserve Bulletin*, (September), pp. 707-725.

Avery, R. and Samolyk, K.A. (2003). "Banks consolidation and small business lending: The role of community lenders." Federal Reserve Board of Governor, Working Paper, pp. 2003-05.

Bostic, R. and Robinson, B. (2003). "Do CRA agreements influence lending patterns?" *Real Estate Economics*, 31, 1, pp. 23-51.

Bostic, R. and Robinson, B. (2005). "What makes Community Reinvestment Act agreements work? A study of lender responses." *Housing Policy Debate*, 16, 3 and 4, pp. 513-545.

Calem, P., Gillen K. and Wachter S. (2004). "The neighborhood distribution of subprime mortgage lending." *Journal of Real Estate Finance and Economics*, 29, 4, pp. 393-410.

Cameron, A. C. and Trivedi, P. (1998). "Regression analysis of count data." *Econometric Society Monographs*, 1998.

Caskey, J. P. (1994). "Bank representation in low-income and minority urban communities." *Urban Affairs Quarterly*, 29, 4, pp. 617-638.

Caskey, J.P. (1991). "Check-cashing outlets in the U.S. financial system." *Economic Review* (Federal Reserve Bank of Kansas City), 76, 6, pp. 53-67.

Chang, A., S. Chaudhuri and J. Jayaratne (1997). "Rational herding and the spatial clustering of bank branches." Federal Reserve Bank of New York, Research Paper Number 9724.

Dahl, D., Evanoff D.D., and Spivey M.F. (2000). "Does the Community Reinvestment Act influence lending? An analysis of changes in bank low-income mortgage activity." Number WP-00-6, Working Paper Series from Federal Reserve Bank of Chicago.

Felici, R. and Pagnini M. (2004). "Distance, bank heterogeneity and entry in local banking markets." Economic Research Department, Bank of Italy, Economic Working Paper number 557.

Hossain, A. R. (2004). "Past, present and future of the Community Reinvestment Act: A historical perspective." Working Paper Series, University of Connecticut.

Immergluck, D. (2004). "Credit to the community: Community reinvestment and fair lending policy in the United States." Armonk, N.Y.: M.E. Sharpe.

Kwan, S. (2004). "Banking consolidation." *Federal Reserve Band of San Francisco Economic Letter*, June 18.

Kiser, E. (2002). "Household switching behavior at depository institutions: Evidence form survey data." *The Antitrust Bulletin*, 42, p. 619.

Kutler, J. (1996). "Stories of the branch's demise have been greatly exaggerated." *American Banker*, 31, (December).

Lang, W. and Nakamura, L. (1993). "A model of redlining." *Journal of Urban Economics*. 33, pp. 223-234.

Li, X., Hossain A. R. and Ross S. L. (2010). "Neighborhood information externalities and the provision of mortgage credit." University of Connecticut, Working Paper No. 2010-10.

Ross S. L. and Yinger J. (2002). "The color of credit: Mortgage discrimination, research methodology, and fair-lending enforcement." MIT press.

Seidman E. and Tescher J. (2004). "From unbanked to homeowner: Improving the supply of financial services for low-income, low-asset customers." Harvard Joint Center for Housing Research, Working Paper Series BABC 04-4.

Appendix 1. Variable Definition

Variable Name	Description	
Dependent Variables		
NOB_{TY}	Total number of branches of all lenders that are within 1-mile distance of tract center.	
NOB_{LTY}	Total number of branches of individual lenders that are within 1-mile distance of tract center.	
Explanatory Variables		
Neighborhood Characteristics	Census Based Variables	HMDA Based Variables[1]
Tract Median Income (in 1000)	Median income of the tract in 1990 census.	Median income of the HMDA applicant pool
Percentage Black	Percentage of Black population in the tract.	Percentage of Black population in the HMDA applicant pool.
Percentage Asian	Percentage of Asian population in the tract.	Percentage of Asian population in the HMDA applicant pool.
Percentage Other	Percentage of Hispanic, American Indian and other population in the tract.	Percentage of Hispanic, American Indian and other population in the HMDA applicant pool.
Percentage Male	Percentage of male population in the tract.	Percentage of male population in the tract.
Percentage Old	Percentage of population in the tract who are 60 years or older.	Percentage of population in the tract who are 60 years or older.
Percentage Occupied	Percentage of occupied housing units in the tract.	Percentage of occupied housing units in the tract.
Percentage Owner Occupied	Percentage of owner occupied housing units in the tract.	Percentage of application in HMDA sample for owner occupied housing units in the tract.
Population Density	Population per sq mile in the tract.	Weighted average population density between 1990 and 2000 census assuming equal population growth.
Lender Characteristics		
Head Distance	Distance between lenders' head office location and center of tracts	
Deposits (in million)	Total dollar amount of deposits received by the lender	
Equity (in million)	Total dollar amount of equities of the lender	
Past CRA ratings (Average of 3 years)	Average of last 3 years of CRA rating	

Appendix 2. Descriptive Statistics

Variable Name	Census Data				HMDA Data			
	Tract Level Analysis		Lender Level Analysis		Tract Level Analysis		Lender Level Analysis	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Dependent Variables								
NOB_T	4.82	6.84	---------	---------	4.54	6.06	---------	---------
NOB_{LT}	---------	---------	0.035	0.268	---------	---------	0.034	0.258
Explanatory Variables: Neighborhood Characteristics								
Tract Median Income	44.24	19.32	44.24	19.32	60.32	30.95	70.34	45.67
Percentage Black	9.08	16.45	9.08	16.45	8.03	17.52	0.08	0.17
Percentage Asian	1.47	1.57	1.48	1.56	1.97	4.7	0.02	0.04
Percentage Other	3.77	8.16	3.77	8.17	6.32	13.43	0.06	0.13
Percentage Male	48.53	4.52	48.52	4.52	50.99	8.28	0.70	0.14
Percentage Old	13.73	6.33	13.73	6.33	---------	---------	---------	---------
Percentage Occupied	93.05	5.87	93.05	5.86	---------	---------	---------	---------
Percentage Owner Occupied	64.09	26.58	64.09	26.58	61.54	48.45	0.92	0.13
Population Density	3826.72	4936.21	3826.72	4935.84	3705.21	4739.20	3705.21	4739.20
Explanatory Variables: Lender Characteristics								
Head Distance	---------	---------	35.66	20.02	---------	---------	35.72	20.05
Deposits	---------	---------	487.57	953.54	---------	---------	487.53	953.35
Equity	---------	---------	53.82	107.94	---------	---------	53.79	10.79
Past CRA ratings	---------	---------	1.93	0.38	---------	---------	1.93	0.38
Number of Years	8		8		8		8	
Number of Tract	823		823		798		798	
Number of Lender	----		108		----		108	
Number of Obser-vation	6584		466,641		6322		448,312	