WHAT CAUSES RACIALLY SEGREGATED RESIDENTIAL COMMUNITIES?

TESTING AGENT-BASED AND INSTITUTIONAL EXPLANATIONS FOR THE

LOCATIONAL DECISIONS OF HOMEBUYERS

By

Michael Henry Nelson

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Community Research and Action

December, 2012

Nashville, Tennessee

Approved:

Professor Paul Speer

Professor Beth Shin

Professor Douglas Perkins

Professor David Padgett

To my wonderful parents, Diana and Bill.

This is as much yours as it is mine.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF MAPS

CHAPTER I

INTRODUCTION

American cities are still significantly racially segregated. But why? Is it because economic differences between racial groups lead to being able to afford different neighborhoods? Is it because people like to live with people of similar racial and ethnic backgrounds? Is it because racial biases in the housing market keep many minorities in underserved minority neighborhoods? Some combination of two, or perhaps all three? This paper explores these questions and aims to assess the influence of each possible explanation.

This dissertation focuses on the persistence of racial residential segregation in Nashville, Tennessee. The questions mentioned above speak to the three standard explanations of racial residential segregation: those that attribute segregation to socioeconomic differences, those that attribute it to ethnocentric neighborhood preferences, and those that attribute it to racial biases in the housing market. Historically, empirical research has found little support for socioeconomic explanations of segregation, but some recent research has suggested that it is of rising importance (Brown & Chung, 2006; Brown & Chung, 2008; Chung & Brown, 2007). Housing market simulation models, based on the theoretical models of economist Thomas Schelling (1971), present a strong case that ethnocentric neighborhood preferences provide a sufficient explanation for contemporary levels of segregation. However, empirical

validation has been elusive. There is no doubt that racial housing discrimination still exists in contemporary American cities; however, several contemporary scholars have doubted the causal role of discrimination in maintaining segregation (e.g., Macy & Rijt, 2006).

The current study addresses this research with three linked empirical analyses. The first tests the ethnocentric preference model of segregation by focusing on home loan applications by race and their connection to neighborhood racial composition. This is done to assess whether the predictions of Schelling's theoretical model of segregation describe homeowner behavior in Nashville accurately. The next two analyses take different approaches to understanding one aspect of bias in housing markets, namely behavior by lending institutions that may promote segregation. One of these examines the role of lending institutions and whether their loan approval/denial decisions have an impact on segregation at the level of census tracts. The other looks at whether there is evidence that lenders show racial bias in their decisions to make loans to particular applicants, and whether lenders are more likely to deny loans for minority households who attempt to move into white neighborhoods. This type of discrimination, as opposed to "redlining" or discrimination on the basis of the race of applicant alone, could directly maintain contemporary high levels of segregation. All analyses control for economic factors. There is no question that as long as wealth is unevenly distributed by race, such factors contribute to racial segregation; the question of this study is whether explanations based on financial factors are sufficient, or whether individual preferences and decisions by lending institutions make additional contributions.

Thus this study analyzes homebuyer loan applications to assess the intersection between socioeconomic resources, consumer-level decisions (preferences), and those of the institutions that serve as gatekeepers to home mortgage financing. By doing so, it seeks to identify whether institutional and/or consumer factors are maintaining segregation and, concomitantly, which types of intervention would be most promising in helping to reduce segregation.

This paper opens with a discussion of recent patterns in segregation and its consequences for minority households in the US. This chapter then provides an overview of the three main explanations of segregation and introduces the three research questions that will be addressed by the proposed study. Finally, it takes a closer look at the relevant literature and research methods of each proposed analyses.

Segregation and its Consequences

Racial residential segregation is an important urban policy issue. American metropolitan areas continue to be racially and ethnically segregated (Logan, 2001). While there are trends toward greater integration, some research suggests that spatial segregation, at least in some form, will be a part of the U.S. landscape for the long-term (Krivo & Kaufman, 1999).

In the case of Nashville, there has been a slight decline in segregation over the past three decades. In 1980 the Nashville metropolitan region had a black-white

segregation index of 0.65[1], meaning that 65 percent of whites would have to move in order for there to be proportional representation of whites and blacks in every neighborhood. Over the next three decades, it fell to 0.61 in 1990, 0.58 in 2000, and 0.55 in 2010. At this level of segregation, 55% of whites would have to move in order for Nashville to be fully integrated.

Segregation has important consequences, the vast majority of which are negative, for minority families and their life opportunities (Acevedo-Garcia, Lochner, Osypuk, & Subramanian, 2003; Bullard, 2007; C. A. Collins & Williams, 1999; Cutler & Glaeser, 1997; Cutler, Glaeser, & Vigdor, 2008; Dawkins, Shen, & Sanchez, 2005; Echenique & Fryer, 2007; Flippen, 2004; Grady, 2006; Grady & McLafferty, 2007; Kim, 2000, 2003; Krivo & Kaufman, 2004; Macpherson & Sirmans, 2001; Pulido, 2000; Quercia, McCarthy, Ryznar, & Talen, 2000). In one particularly thorough study, Cutler and Glaeser (1997) show that, for the average city, a one standard deviation reduction in segregation would reduce the black-white differential in educational performance, teen pregnancy, and employment by one-third. The deleterious impact of segregation makes understanding its causes an important and contentious area of research.

Explanations of Racial Residential Segregation

As mentioned above, there are three broad explanations of contemporary racial segregation: those that focus on socioeconomic differences, those that emphasize

---

[1] This was calculated using the index of dissimilarity, a common measure of segregation.

ethnocentric neighborhood preferences, and those that highlight the racial biases at work in the housing market. Comprehensive reviews of the literature on these theories can be found elsewhere (e.g., Dawkins, 2004). Here, I summarize the key points and findings to introduce my research questions.

Socioeconomic resources delimit housing options. If a household has enough money, it can afford to purchase any house it desires; if it doesn't, some houses and neighborhoods may be out of financial reach. Thus some scholars emphasize how income, wealth and education differences between racial groups undoubtedly increase the potential for segregation. Accordingly, there is some evidence that the relationships between an individual's SES and neighborhood characteristics may have become stronger in recent years (Brown & Chung, 2006; Brown & Chung, 2008; Chung & Brown, 2007).

Historically, however, empirical research on locational attainment – that is, studies which look at the relationship between household characteristics and neighborhood racial composition and median income – consistently find that householder race is a far better predictor of neighborhood characteristics than their socioeconomic status (Alba & Logan, 1992; Alba, Logan, & Stults, 2000; Charles, 2003; Crowder, South, & Chavez, 2006; Dawkins, 2004; Jones, 2008; Logan, Alba, & Leung, 1996; Pattillo, 2005; Rosenbaum & Friedman, 2007; South, Crowder, & Chavez, 2005; South & Crowder, 1998)..

Given that household socioeconomic status has been a weak predictor of neighborhood racial composition, other scholars point to the importance of neighborhood preferences in shaping metropolitan housing patterns by race. There is evidence to

suggest that different racial and ethnic groups prefer to live in neighborhoods in which they are the majority (Clark, 2002; Clark & Fossett, 2008). Since it is impossible for all groups to live in neighborhoods in which they are the majority and for neighborhoods to be simultaneously fully integrated, it is argued that segregation is the natural outcome of an unbiased housing market. Computer-based models, based on the theoretical work of Thomas Schelling (Schelling, 1971), have provided substantiation of these claims under a wide variety of assumptions (Clark & Fossett, 2008; Fossett, 2006a, 2011; Fossett & Dietrich, 2009; Fossett & Waren, 2005; Laurie & Jaggi, 2003; Wasserman & Yohe, 2001; Zhang, 2004). Some scholars, however, are skeptical of the results generated by these computer models (e.g., Goering, 2006). Further, empirical studies have found only mixed support for preference-based theories of segregation (e.g., Adelman, 2005; Freeman, 2000). Proponents of preference-based models of neighborhood segregation counter that their critics have "rarely engaged Schelling's celebrated theoretical treatments directly and certainly have not refuted them" (Fossett, 2006b, p. 295).

The lack of convincing empirical evidence for both socioeconomic explanations and neighborhood preference-based explanations has led a third group of scholars to argue that the racial biases of institutions operating in the housing market are an important factor in maintaining contemporary segregation. Evidence of racial bias has been found at nearly all aspects of the housing search, beginning with the way different houses are advertised based on the neighborhood's racial composition (M. Collins & Galster, 1995), the way real estate agents and property managers interact with potential clients and the information they are provided (Fischer & Massey, 2004; Massey & Lundy, 2001; Turner & Ross, 2005), the prices quoted to prospective residents (Turner & Ross,

2005), whether homes are available (Turner & Ross, 2005), and the decision of lenders (Holloway & Wyly, 2001; Ross & Yinger, 2002) and home insurance providers (Squires & Kubrin, 2006). While there exists much evidence on the presence of racial biases in the housing market, there is debate about its relative importance in maintaining segregation, especially given the decline of overt forms of discrimination over recent decades (Brown & Chung, 2008; Turner & Ross, 2005).

Research Questions

The current study enters into this discussion about the relative importance of economics, preferences, and racial bias in maintaining racial residential segregation by asking three linked questions. The first question focuses on preference-based theories of racial segregation and responds to Fossett's call for a direct assessment of the effectiveness of Schelling's preference-based theory. By quantifying the effect of changes in neighborhood racial composition on the racial characteristics of those that attempt to move into that neighborhood, the first question asks: do potential homeowners respond to neighborhood characteristics in a way that is consistent with preference-based models of segregation?

The second question asks: To what extent do the decisions of mortgage lenders reshape the metropolitan distribution of households by race? This analysis will quantify the impact of lending decisions on the residential segregation of neighborhoods. Whereas question one looks at where households apply for loans, question two examines where

households are approved for loans – a question with very little empirical exploration to date. Lending institutions have the ability to reshape the neighborhood distribution of homeowners by race by virtue of their loan denial/approval decisions. This analysis, at the census tract level, quantifies the aggregate impact of lending decisions of the racial composition of in-movers.

The third question builds on the first two analyses by looking at individual loan applications (as opposed to census tracts) and asking: Is there evidence of discrimination in lending decisions and does this lead to greater segregation?   While question two simply investigates whether there may be factors other than preferences (i.e,, lending decisions) playing a role in segregation, question three asks whether there is evidence of discrimination in lending decisions and whether this, specifically, is contributing to racial residential segregation. There are three main types of racial discrimination in lending suggested by Holloway and colleagues (Holloway, 1998; Holloway & Wyly, 2001): (1) discrimination based on the race of applicant; (2) discrimination based on the neighborhood racial composition ("redlining"); and (3) discrimination based on an interaction of race of applicant and the racial composition of the neighborhood ("geographically-contingent lending"). While consistent discrimination against minority households or minority neighborhoods may not have much impact on segregation levels, geographically contingent lending provides a direct mechanism to maintain and expand current levels of segregation.

For each of these questions I will control for the impact of socioeconomic status and assess whether economic factors may explain any differences in outcomes by race. The table below (Table 1.1) describes the key aspects of each question. The table

highlights whether the analysis focuses on would-be homebuyers (consumers) or lenders (institutions), contains a brief summary of the question, the level of analysis (whether the analysis focuses on census tract level patterns or on individual applicants), the key phenomenon of interest (whether it is primarily segregation or discrimination), the dependent variable, key independent variables, and type of model. In the subsequent chapters, I take a more in-depth look of the literature and analyses pertaining to each of these questions.

Table 1.1 Research Questions

	Focus	Question	Level of Analysis	Phenomenon	Dependent Variable	Key Independent Variables	Type of Model
Analysis 1	Consumer Decisions	Do homebuyers in Nashville apply for home loans in a way that is consistent with preference models of segregation?	Neighborhood (tract) – How does the racial composition of applicants relate to neighborhood racial composition	Segregation (consumer level decisions)	Percentage of applicants who are white (by tract, by year).	Estimated tract percentage white, estimated tract SES, applicant median income, applicant median loan amount.	2-level spatial growth model
Analysis 2	Lending decisions	Do the decisions of lenders reshape the distribution of homebuyers by race?	Neighborhood (tract) – How do loan approval decisions by race relate to neighborhood racial composition	Segregation (institutional level decisions)	Difference between % of applicants who are white and % of approved applicants who are white (by tract, by year)	Tract percentage white, tract SES, median white income, median minority income, ratio of median white income to median minority income	2-level spatial growth model
Analysis 3	Lending decisions	Do the decisions of lenders provide evidence of racial bias?	Loan Application – how do applicant and neighborhood characteristics relate to the decision to approve or deny a loan	Discrimination (institutional level decisions)	Loan approval decision (1 = approved, 0 = denied).	Applicant: race, income, loan amount, credit history instrument. Tract: percentage white, SES.	2-level spatial logistic model

CHAPTER II

SCHELLING'S PREFERENCE-BASED THEORY OF RACIAL RESIDENTIAL

SEGREGATION, AGENT-BASED MODELING, AND EMPIRICAL RESEARCH

The preference model of residential segregation, developed by Schelling, holds

that individual households have their own preferred, 'ideal' neighborhood of residence

with regard to racial composition. When a household's neighborhood deviates too far

from their preference they are likely to move, or if in the market to purchase a home, to

look elsewhere for a home. The preference model asserts that patterns of racial residential

segregation may come about from the aggregate preferences of households, rather than

institutional racial biases (e.g., through racial steering by real estate agents or lending

discrimination by banks). Heretofore, the strongest support for Schelling's preference

model of segregation has largely come from theoretical models rather than empirical

data. In this section I will (1) outline Schelling's 3-part theoretical analysis, (2) examine

agent-based modeling that use computer simulations, and (3) look at the most relevant

empirical findings.

Part 1: Simple Line

Schelling (1971) builds his theoretical analysis in three parts. First he simulates segregative tendencies on a one-dimensional line, he then looks at these tendencies in two dimensions by using a chess-board type simulation of a city, and finally he looks at what happens when there are within group differences in neighborhood preferences.

Schelling (1971) begins his analysis simply. The first part of his analysis involves a line made up of a random sequence of plus-signs ("stars" in his paper) and zeros:

0+000++0+00++00+++0++0++00++00++00++0+0+00+++0++00000+++000+00++0+0++0

Figure 2.1 Simple line of random zeroes and pluses, from Schelling (1971, p. 149)

He then models what would happen if every plus wanted at least half of its neighboring symbols to be pluses, and every zero wanted at least half of its neighboring symbols to be zeros (in this basic model, neighbors consist of 4 symbols on either side). Schelling assumes that each plus and each zero does not care how many like neighbors they have, as long as it is above this threshold. He begins by going along the line from left to right and moving any plus or zero that currently does not have half its neighbors (4 of 8) as the same symbol to the nearest point along the line where this condition would be satisfied (the dots in Figures 2.1 and 2.2 denote which symbols are currently unsatisfied). He

12

concludes when every symbol is satisfied with its neighborhood composition. It takes two rounds to make every household satisfied. After the first round there are eight symbols that have become discontented due to the movement of symbols during the first round. After the second round, all symbols are content:

00000000++++0+++++++++0000++000+0+0+++0+++++++++00000000000000000+++++++

00000000++++++++++++++0000000000++++++++++++++00000000000000000++++++
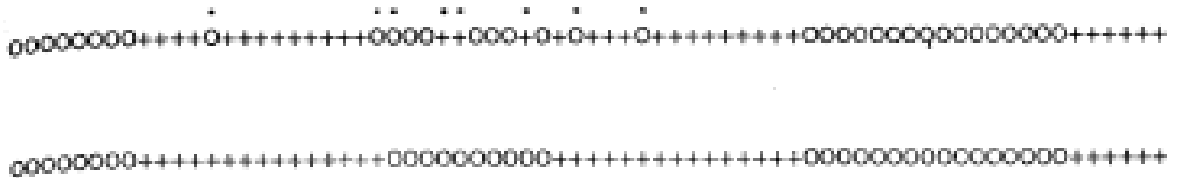
Figure 2.2. First line shows dispersion of symbols after one round of movement (symbols with dots have become unsatisfied), second line shows symbols after two rounds of moves (all symbols satisfied). From Schelling (1971, p. 151)

While simple alternating groups of two (two pluses, followed by two zeros, followed by two pluses, etc.) would satisfy all symbols, he finds that the average group size (i.e., the number of consecutive similar symbols) is 14. This ends up with almost half of all the symbols having no neighbors that are not the same symbol (i.e., nearly half of all zeros do not have pluses within 4 symbols either side of them). Thus, in this very simplistic model, slight ethnocentric (or, in this case, symbol-centric) preferences (a desired ratio of 5:4) tend to yield highly segregated outcomes (a ratio of 5:1).
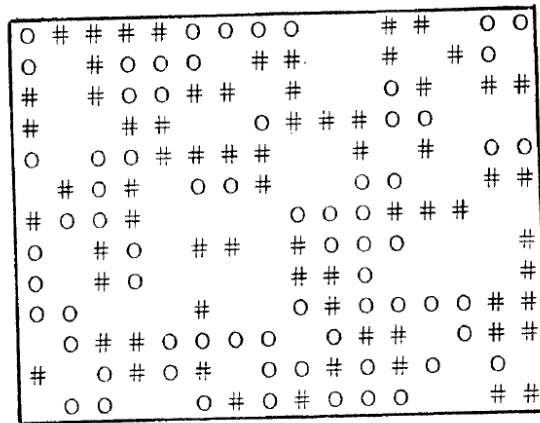
Figure 2.3 Random assortment of hashes and zeroes (Schelling, 1971, p.155)

Part 2: Checkerboard

In part 2, he builds on his analysis by using a 13 by 16 checkerboard (208 individual squares – see figures 2.3 and 2.4) rather than a one-dimensional line and finds that under similar assumptions (this time using each of the eight squares immediately adjacent to the symbol as their neighborhood) of slight ethnocentric preferences (once again, a minimum of 50% same race neighbors) result in high levels of segregation (when they are equal numbers of stars and zeros, each symbol has 80% of their neighbors sharing the same symbol as them), even though no single individual wants to live in a highly segregated neighborhood. Different aspects of the model are then varied including the neighborhood preferences of each group and their relative proportions in the model. Key findings include:

14
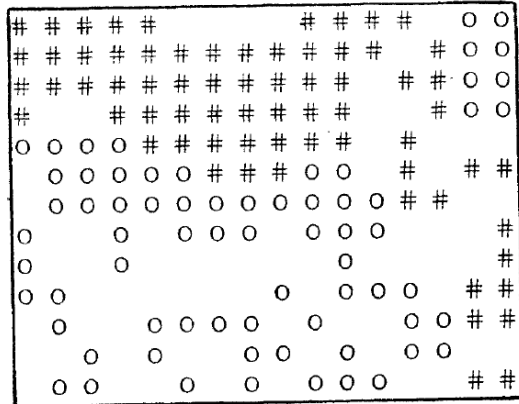
Figure 2.4 After symbols have been moved to areas where their preferences are satisfied (Schelling, 1971, p.157)

- When groups are equal in number and the preference for like neighbors is reduced from one-half to one-third, little segregation results.

- When there are more members of one group than the other, greater segregation results.

- Larger "neighborhoods" (i.e., including more than just the 8 surrounding squares) result in lower levels of segregation.

Part 3: Varying Preferences

After the two-dimensional analysis, in part 3 he focuses on segregative patterns when neighborhoods are fixed in space (everyone within a neighborhood considers the same group of households as their neighbors, rather than just those that are relatively close to them). On top of this, he adds complexity by assuming residential racial preferences were distributed differently across the population (i.e., not all whites have the same

preferences). Some may consider this model as one that better reflects the reality of housing decisions. In this series of analyses he begins with a model that assumes a 2:1 ratio of whites to blacks in the population. He then assumes that the median person in each group prefers a neighborhood racial ratio of one white household to one black household (therefore, half of the whites, and half of the blacks, prefer a neighborhood in which they are in the minority, and the other half of prefer a neighborhood in which they are the majority). In this scenario he finds that while the average person wants to live in an integrated neighborhood, there are only two stable neighborhood states: all white neighborhoods and all black neighborhoods. A variety of modifications to this model leads to the following conclusions.

- In the case of more diversity-tolerant households (median household can tolerate a ratio of 1:2.5 same to other households) and equal numbers of whites and blacks (100 each) there can be a stable state in which 80% of the households live in truly mixed (50-50) neighborhoods.
- However, if this model is changed so that the ratio of whites to blacks is 2:1 (rather than 1:1) then, once again, the only stable neighborhoods would be completely black neighborhoods and completely white neighborhoods.

Neighborhood Tipping

The general pattern of results from Schelling's increasingly complex models suggests that segregation is the "natural" state of affairs when there is a tendency for groups to prefer to not be in a minority. This is more likely to be the case when there is unequal numbers in each group (as is the case in most American cities) and the smaller

the area that is defined as a neighborhood. Segregation is likely to result under these circumstances whether neighborhoods are considered to be relative spaces (within a certain distance) or absolute (hard geographic boundaries dividing one neighborhood from another). Even when there is variation among the residential preferences of each group segregation will still result, as the least tolerant will tend to start a tipping process.

Schelling's theory of tipping begins with the assumption that households have threshold levels of racial composition that shape their decisions whether to move to or stay in a neighborhood. Additionally it assumes that these threshold levels vary from household to household. As each household makes a decision to move into or leave a neighborhood, they 1) potentially change the composition of that neighborhood, 2) affect the move or stay decisions of their neighbors, and 3) affect the desirability of the neighborhood to potential in-movers. Thus, there will be critical thresholds where, once crossed, a neighborhood's in-movers will more likely be out-group members than in-group members. Once these thresholds are crossed, a neighborhood is likely to "tip". In this Schelling finds a corollary that could potentially explain today's level of segregation, while nobody may want highly segregated cities, a slight preference to be not in a minority leads to unstable mixed neighborhoods that tend to tip into highly homogenous spaces.

Computer Agent-based Modeling

While Schelling's explorations using checkerboards and symbols may be instructive in many ways, his analysis does not come close to resembling the dynamics of actual cities. Recent advances in computing power and programming have made it possible to run Schelling-type simulations that take into account a whole host of parameters deemed pertinent to the study of neighborhood racial patterning. The most advanced example of this is Fossett's SimSeg program that was examined in a 2006 special issue of the Journal of Mathematical Sociology (on the 35th anniversary of that same journal's publication of the seminal Schelling paper discussed above).

Fossett (2006) described his SimSeg program working as follows:

1. The program creates a virtual city composed of bounded neighborhoods arranged in a grid. Each neighborhood contains a sub grid of a fixed number of houses.

2. The program creates housing units of varying quality and fills each subgrid with these houses.

3. The program creates a population of households with each household having an ethnic status and socioeconomic status (simplified to a number representing household income). Each household also has a specific preference for housing quality, neighborhood socioeconomic status, and neighborhood ethnic composition.

4. These households are randomly assigned to housing units, but only to housing units they can afford.

5. The program calculates the level of segregation that exists at the beginning and these measures are stored as a baseline measure.

6. The simulation commences by selecting households at random and these households conduct a limited search to see if they can find a housing unit that better meets their set of housing quality, neighborhood status, and ethnic composition preferences.

7. This repeats for a designated number of cycles and takes periodic measures of segregation to measure change from the baseline measurement.

Fossett begins his 2006 paper by using his program to examine whether SES differences could explain segregation. He then turns to five more models that analyze what happens to levels of segregation under different neighborhood racial composition preferences. The population demographics for all of his analyses consist of 60% white households, 20% black households and 20% Hispanic households.

Model 1: Socioeconomic Status Differences

In his examination of whether socioeconomic status differences could explain segregation, Fossett tests what would happen if households cared only about housing quality and neighborhood status and not ethnic composition. To test whether socioeconomic status ("income") differences would lead to segregation, Fossett sets housing value to vary with distance from the center of the "city" (so the neighborhoods furthest away have the highest average values) and then sets the median income for Hispanic households at 75% of the white median income, and black households at 63% of white median income. At the outset, segregation is low and after 30 cycles it remains

low, with an average normed index of dissimilarity (ID) of 0.14^2. These results lead Fossett to suggest that economic segregation may not provide a good explanation for racial segregation.

Models 2 and 3: In-Group Preferences in Line with Survey Data

In the next series of models he tests what would happen under varying levels of ethnocentric preferences. Fossett begins with two models that he argues have preference schedules that are in line with those found in survey research. The only difference between Models 2 and 3 is the way households look at their neighborhoods. Households in Model 2 consider neighborhoods to be fixed in space (all households in a neighborhood grid consider each other neighbors), while households in Model 3 take into consideration only nearby neighbors within a fixed distance (which may cross the "neighborhood grid" boundaries).

For these models, Fossett sets median preference at 90% same race/ethnicity for whites and 50% same race/ethnicity for blacks and Hispanics. This, Fossett suggests, is consistent with what is found in survey research. Fossett sets minority preference for out-group neighborhood composition to be 30% and both black and Hispanic prefer their out-group neighbors to be white (this is also in line with survey research).The out-group preferences are subordinate to their in-group preferences, however. While whites don't

---

[2] The index of dissimilarity (ID) provides a measure of what proportion of households would have to move in order for there to be proportional representation of each racial/ethnic group in every neighborhood. Fossett runs each model several times and reports the average ID that results. In this case, an average of 14% of given group's households would have to move – a number much lower than most U.S. cities.

have preferences for a minimum number of out-group neighbors, they prefer

neighborhoods with Hispanic households to black households[3].

Both models result in high levels of segregation.   Model 2 results in an average

index of similarity of 0.84 (84% of a given group's households would have to move).

Model 3, which was similar to Schelling's checkerboard analysis, results in an average

ID of 0.81 (although he notes that the patterning is slightly different from model 2).

Model 4: Proportional Preferences

Fossett then relaxes the ethnocentric preferences of each group. First, he models

what would occur when neighborhood ethnic preferences are set to be simply in line with

each group's representation in the population: median white preference is for a 60%

white neighborhood, Hispanic and black households' median preference is for 20% in-

group neighbors (the out-group preference of 30% is dropped in this analysis). Once

again, relatively high levels of segregation result (ID = 0.60). This is a particularly

striking result as this set of preferences would seem to align with what should result in an

index of dissimilarity of zero: every neighborhood having proportional representation of

each group.

Models 5 and 6: Stronger Out-Group Preferences

For Models 5 and 6, Fossett assesses what happens when out-group preferences

were strengthened, while maintaining in-group preferences at the same proportional level

---

[3] It is worth keeping in mind that neighborhood ethnic preferences are just one of three difference
categories of preference, and therefore play only a partial role in determining where households choose to
live (the others remain housing quality and neighborhood status) and households will move to an area with
an ethnic mix that is not in line with their preferences if that unit meets the sum total of their three types of
preferences better than any other option.

as they were in Model 4. He does this in two ways. First, Fossett sets the minority (black and Hispanic) out-group preference back to 30% much as it was in the earlier analyses. In the second analysis, he sets it to 60%. Not surprisingly, these both significantly reduce segregation levels. The first reduces segregation by 14 points (ID = 0.46). The second reduces it an additional eight points to an ID of 0.38.

Model 7: Final Model and Conclusion

While these lower levels of segregation result from plausible preferences, Fossett notes that these preference schedules depart significantly from what is suggested by survey data, and returning white preference to 90% white neighborhoods and moving minorities' preference for in-group neighbors to 25% rather than 20% results in a segregation index of 0.60. Fossett concludes that the neighborhood preferences found in survey research are more than adequate to create highly segregated cities.

Thus, if one imagines actual households as rational agents with preference schedules that fall along the lines articulated by Schelling, Fossett and others, we would predict that the outcome of unfettered household locational decisions would be high levels of segregation not unlike what is found in Nashville, Tennessee or any other major city in country.

Commentary on Fossett's Models

Fossett's article received much praise from the scholars commenting on his paper in the special issue of JMS. For example, Clark (2006) commented: "Fossett's new research… is arguably the most important advance in studies of residential segregation in the past decade…. The paper shows clearly that preferences do matter and that residential

preferences and their underlying social dynamics have the capacity to generate high levels of ethnic segregation." (Clark, 2006, p. 319) Similarly, St. John (2006) felt that the technical sophistication of Fossett's analysis, allowing for the complexities and nuances unmodeled in previous research, made for a compelling case that preferences are maintaining the high levels of segregation that are seen in most American cities. However, a third paper argues that Fossett's conclusions do not go far enough (Macy & Rijt, 2006). Macy and Rijt argue that Fossett's modeling program shows that, not only can preferences account for current levels of segregation, but housing discrimination, by itself, could not account for high levels of segregation. That is, discrimination, without ethnocentric preferences, is not a necessary - nor sufficient - cause of segregation. More recent articles by Fossett and coauthors using SimSeg find continued support for Schelling's basic conclusions on preferences (Clark & Fossett, 2008; Fossett, 2011; Fossett & Dietrich, 2009).

Only one of the four commenting papers (Goering, 2006) was strongly skeptical of Fossett's methods and conclusions. Goering critiqued Fossett for his reductionist approach to examining segregation. Goering suggested that the very things not accounted for by the model are some of the most crucial in understanding the patterns residential racial segregation. He cited the uneven distribution of amenities, capital investment, and disinvestment as crucial aspects of housing patterning. He argued that an analysis without these factors can provide little insight into actual segregative dynamics. Thus, Goering believes that further empirical research, as opposed to computer simulations, is likely to be a more productive research direction. I now turn to the most pertinent of the few empirical evaluations of preference models of segregation.

Empirical Research

The analysis of Card et al. (2007) provides empirical support for Schelling-type neighborhood tipping. They find the neighborhood racial preferences of whites to be crucial in the dynamic process of tipping, and find little support for socioeconomic concerns as the driver of racial tipping. Card et al. use a regression discontinuity design to look for non-linear change in neighborhood racial composition from one decennial census to the next. Using census data at the tract level for 61 metropolitan areas from 1970 to 2000, the authors find neighborhood tipping points in most cities. While considerable variation in tipping points across metropolitan areas exist, the results indicate that, on average, a census tract would tip when it reached approximately 13% black. Tipping was most pronounced in the 1970s and 1980s but declined during the 1990s. They also found a negative correlation between each metropolitan area's prejudice level[4] and neighborhood racial tipping point: Areas with higher prejudice levels had neighborhoods that would tip sooner (with fewer black neighbors) than areas with lower prejudice levels. On the whole, the instability of neighborhood racial composition around the tipping points provides support for Schelling's tipping theory.

Even though, in general, Card et al.'s (2007) empirical research seems to support preference models of segregation, some of their findings add considerable complexity to the understanding of segregative dynamics articulated by preference theorists. First, their

---

[4] The authors use the metropolitan average of white responses to questions on racial bias from the General Social Survey.

findings demonstrate that neighborhood tipping has seemed to have weakened across time. This is not necessarily in line with what would be predicted by a strong version of the conclusions of Fossett – as long as no group wants to be in the minority, we would expect high levels of segregation. Second, that prejudice levels are associated with tipping points suggests that there is something more at play than simple homophily. If segregation is associated with prejudice, rather than just homophily, one has to wonder whether the institutions that work within the housing market are also imbued with some prejudice.

To explore this question, I now turn to two studies that, instead of looking at aggregate change in neighborhoods between decennial census years, examine the stated preferences of households and compare these preferences to their current neighborhood racial composition. First, Freeman (2000) focuses on the differences between minority groups in their segregation from white households. That is, he compares the neighborhood exposure to whites of black, Latino, and Asian households using data collected from the Multi-City Study of Urban Inequality (MCSUI). Freeman asks: do minority groups (i.e., black, Asian, and Latino households) have equivalent proportions of whites in their neighborhood given similar neighborhood preferences and socioeconomic status. Freeman finds that neighborhood residential preferences are a strong predictor of neighborhood racial composition, but that black households with similar socioeconomic, lifecycle, and preference characteristics are much less likely to be sharing a neighborhood with white households when compared to both Latinos and Asians. That is, preferences predict neighborhood racial composition, but only to a limited degree. Preferences matter, but so does the minority group which one is a

member. Freeman suggests that the greater segregation of blacks from whites may be the outcome of discriminatory institutions at work in the housing market.

Similarly, Adelman (2005), also using MCSUI data, compared middle-income white households with middle-income black households and found that, while preferences are important predictor of neighborhood composition, white households with preferences for integration tend to live in neighborhoods that are 85% white while black households with similar preferences tend to live in neighborhoods that are 30% white. Adelman suggests that powerful social forces, like racial steering and lending discrimination, are limiting the opportunities of middle class blacks to move into integrated neighborhoods.

Both Adelman and Freeman conclude that their results indicate that a significant institutional bias exists against blacks, which, in turn, maintains racial residential segregation. However, Schelling's models predict that racial change can happen quite rapidly, and this is borne out in Card's analysis. Therefore, it could be that when households moved into their neighborhoods they were more diverse than they were at the time of the study. Also, given the lack of integrated neighborhoods in most cities, it could be that the type of neighborhoods black households would like to live in do not exist (either due racial tipping as described by Schelling and shown in Card, or because of previous housing discrimination). However, without more empirical research on contemporary housing decisions it is difficult to assess the relative strength of these competing explanations. Thus, we have a situation in which there is a compelling argument that ethnocentric neighborhood preferences can account for contemporary

levels of segregation, but a distinct lack of empirical research that conclusively supports

the tenets of these preference theorists.

CHAPTER III

EMPIRICALLY EVALUATING PREFERENCE MODELS OF SEGREGATION

The first research question attempts to evaluate Schelling's and Fossett's models by testing the strength of preference theory in predicting homebuyer decisions in Nashville over an eleven year period (2000-2010). In order to do this, I analyze the relationship between the racial composition of potential movers into a neighborhood and that neighborhood's existing racial composition. By using Census and Home Mortgage Disclosure Act (HMDA) data, this analysis assesses the annual racial composition of mortgage applicants in each neighborhood, and by extrapolating from decennial Census data it estimates the actual neighborhood racial composition on a yearly basis. This shows whether the racial composition of applicants changes as the neighborhood racial composition changes, exploring the central thesis of the preference-based models of segregation.

The hypothesis to be tested is as follows:

As a tract becomes whiter in its racial composition, the proportion of applicants applying to that neighborhood will become whiter too. Similarly, as a tract becomes less white, the proportion of applicants applying to that neighborhood will become less white. This will be the case even after controlling for socioeconomic changes in the neighborhood as well as the socioeconomic characteristics of applicants.

28

This question asks whether mortgage applicants behave, in aggregate, in a way consistent with the processes modeled by Fossett. By translating Fossett's theoretical approach to an empirical analysis, the present study highlights a key limitation to extrapolations based on preference based theoretical models. It does not account for any processes that may bias the racial composition of applicants prior to the application process[5]. The Housing Discrimination Study (Turner, Ross, Galster, & Yinger, 2002) showed that significant levels of discrimination against minorities were occurring at every stage of the home search and, in particular, racial steering by real estate agents had increased between 1989 and 2000 (Galster & Godfrey, 2005). There is also evidence that potential applicants receive differential treatment by lenders based on race before applicants submit loan applications (Ross, Turner, Godfrey, & Smith, 2008). Thus, the question posed in this chapter assesses whether a Schelling-type pattern of applications is occurring. It cannot assess the causes of this pattern. Lack of evidence for a Schelling type pattern will, however, cast strong doubt upon the value of his and Fossett's simulations in explaining racial residential patterning.

To carry out this analysis, a spatial 2-level hierarchical linear model (SHLM; Savitz & Raudenbush, 2009) will be employed to assess how the racial composition of loan applicants changes as the racial composition of the neighborhood changes through this eleven-year period. The data have a natural nested structure, with each census tract having up to 11 observations (one for each year), and, therefore, are well suited to hierarchical modeling. The data are also spatial, with each census tract closer to some and further away from other tracts. Ignoring this spatial factor would lead to potentially

---

[5] Chapters 4 and 5 examine the role of institutional constraints in the housing market that may lead to segregation.

misspecified models and downwardly biased standard errors (Fotheringham, Brunsdon, & Charlton, 2000). Using SHLM addresses spatial dependence by explicitly modeling the spatial correlation in the error terms of adjacent census tracts (Savitz & Raudenbush, 2009). In SHLMs, the neighborhood level error is modeled as having two parts: the random component and the spatially dependent component. The spatially dependent component has three aspects: the spatial weight matrix (which delimits the proximity of one neighborhood to another[6]), the spatial correlation parameter (zero if there is no spatial dependence in the error term and positive [negative] if closer neighborhoods have similar [dissimilar] error terms), and a vector of random spatially autoregressive effects. More detail on the actual equations is provided below.

The dependent variable is the percent of applicants that are white for each census tract in each year. The key independent variables are the year of application and the neighborhood percentage white (NPW) and how it has changed over time (CPW). The control variables from the decennial Census data include the estimated neighborhood socioeconomic status in 2000 (NS; an index variable based on neighborhood median income, percent of persons with a college degree, percent of homeowners) and the change in neighborhood socioeconomic status (CNSES). The control variables from the annual HMDA data include median loan amount and median applicant income.

---

[6] In these analyses, proximity is defined as contiguous/not contiguous (i.e., the matrix is composed of zeros and ones).

Data

HMDA Data

Home Mortgage Disclosure Act (HMDA) data provides information on borrower

and loan characteristics from a majority of lenders[7]. The area covered by this study

includes Davidson County and its six adjacent counties (see map 3.1). This area was

chosen as it provides comprehensive coverage of Nashville's housing market (almost

88% of the population of the 13 county Nashville Metropolitan Statistical Area), without

extending too far into rural communities that have housing markets that are relatively

independent from dynamics within Nashville's housing market. In 2000, this seven

county area had a population of 1,188,000 people: 77.5% of which identified as non-

Hispanic white, 16% identified as non-Hispanic black, and 3% identified as

Hispanic/Latino.

---

[7] For detailed information on HMDA data and its uses see http://www.ffiec.gov/hmda/history.htm and
Avery, R. B., Brevoort, K. P., & Canner, G. B. (2007). Opportunities and issues in using HMDA data.
Journal of Real Estate Research, 29(4), 351-379.
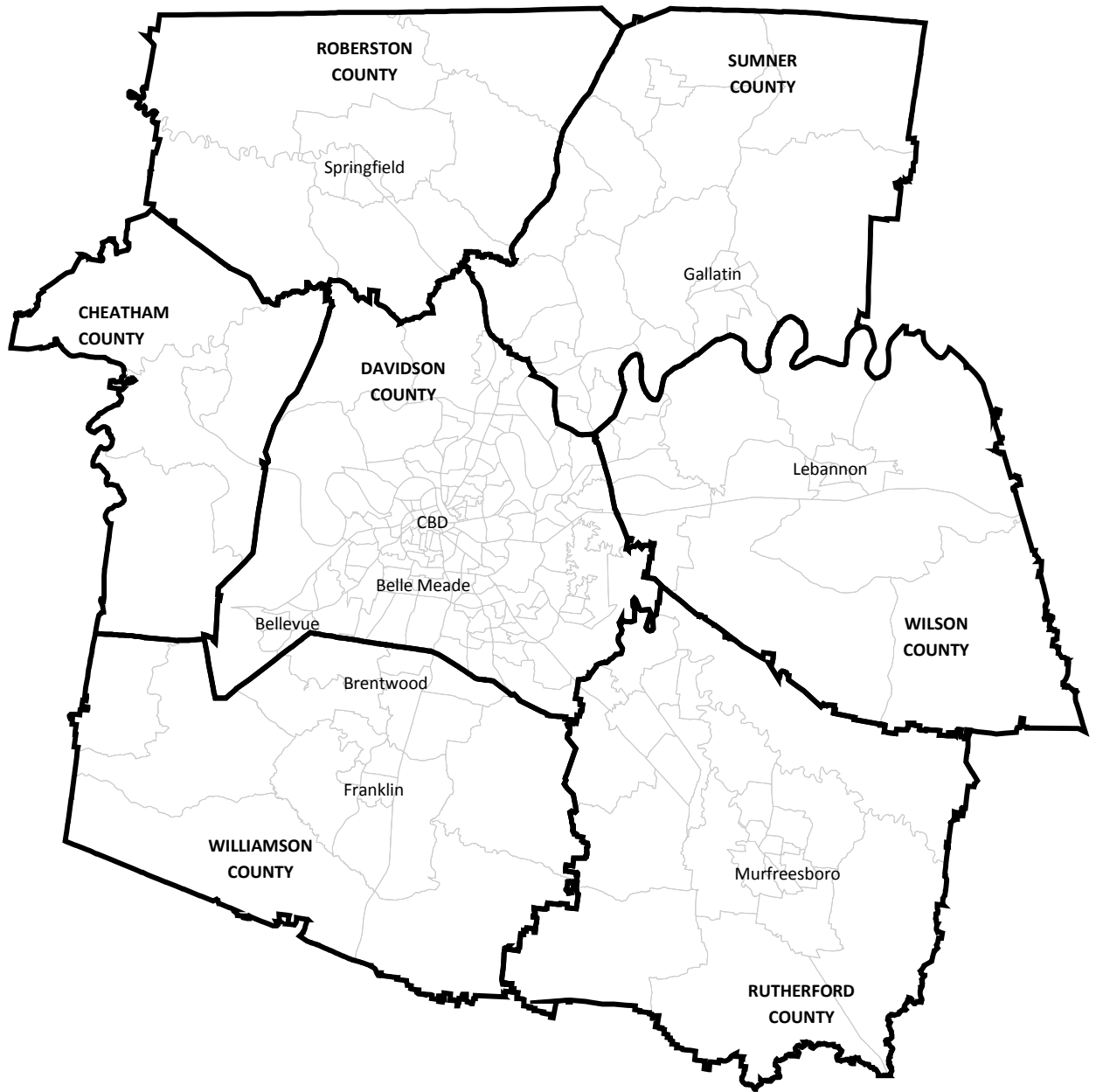
Map 3.1 Nashville Metropolitan Area

From the 1,602,775 loan records in the 2000-2010 in the Nashville MSA HMDA Loan
Application Register, I included the individual loan records where the following were
true [applications remaining]:

1. It was a home purchase loan [n = 657,136]

2. The borrower was going to reside in the home [598,163]

3. It was within a valid census tract [598,066]

4. It was within one of the seven counties [570,879]

5. Applicant race was reported [442,658]

6. Loan amount was between $10,000 and $1,000,000 [439,644].

7. I removed records that did not have incomes between $1 and $1,000,000
   [425,585].

After the data was opened in R statistical software, I performed the following operations:

- I then adjusted the income and loan amount variables by inflation using the
  Bureau of Labor Statistics inflation calculator. All years were normalized to 2009
  dollar amounts.

- I calculated, for each year, each census tract's median loan amount and median
  income, as well as the number of applicants and the number of white applicants.
  This new table of tract characteristics for each year formed the basis for the
  analysis (one can think of these observations as "year-tracts").

- To ensure that the dependent variable was calculated based on a substantial
  number of observations, any year-tract that had fewer than 20 loan applications
  was dropped (144 observations were dropped). This resulted in 2,297
  observations (year-tracts). The average tract had between 8 and 9 years of data.

Census Data

One challenge associated with covering an extended period of time is that census tract definitions change. HMDA data used 1990 census geographies up until 2003/2004, and then used 2000 tract geographies (inspection of 2003 and 2004 data shows that both 1990 and 2000 census geographies were used). In general, the number of census tracts increases with population. Thus, 1990 tracts that have significant population increases tend to split into two or more tracts in 2000 (these new tracts are often joined with portions of adjacent tracts, making simple 1 -> 2 relationships unlikely).

Attaching 2000 data to the 1990 census tracts is a complicated procedure. Thankfully, the Census Bureau provides a table that lists every 1990 tract and provides a link to 2000 census tracts[8]. This file was used to merge the 2000 data with the 1990 tracts, controlling for the proportion of the 2000 census tract population that was within the 1990 census tract. More detail on how this was carried out can be found in the appendix at the end of this chapter.

Once it was possible to calculate neighborhood characteristics for 1990 tracts in 2000, I then calculated the racial composition and neighborhood socioeconomic status (NSES) of each neighborhood for each year. To do this, I used census tract data from the 2000 decennial census as well as the 2005-2009 American Community Survey (ACS). The neighborhood socioeconomic status (NS) variable was created using a principal components analysis of the percentage of residents over 25 that have at least a bachelor degree, the percentage of homeowners, the median value of homes and the median

---

[8] http://www.census.gov/geo/www/relate/rel_tract.html

household income. I removed any census tracts that had missing values on any of these variables (6 tracts from 1990 geographies and 10 tracts from 2000 geographies were removed). I then used the factor loadings of an unrotated principal components analysis as the measure of neighborhood SES (NS). (In 2000, the primary factor accounted for 65% of the variation of these variables. In 2005-2009, it accounted for 71%.) These loadings have an approximate mean of zero and standard deviation of 1. See Figure 3.1 for a description of the results of each of the analyses. The maps show the factor loadings by tract, split into quintiles.
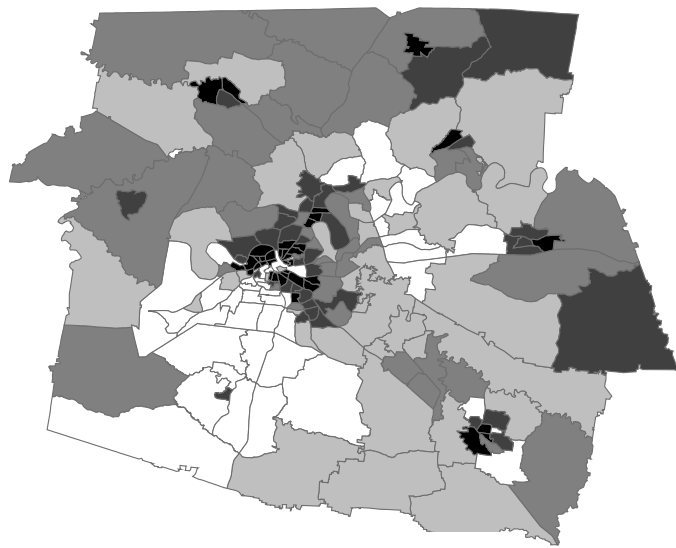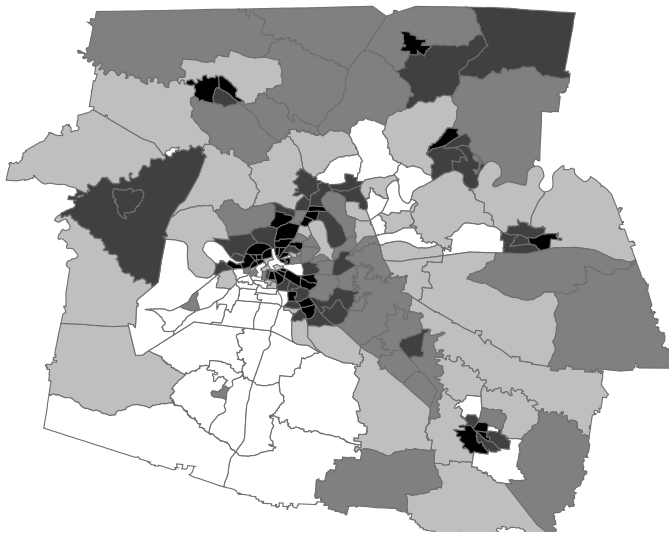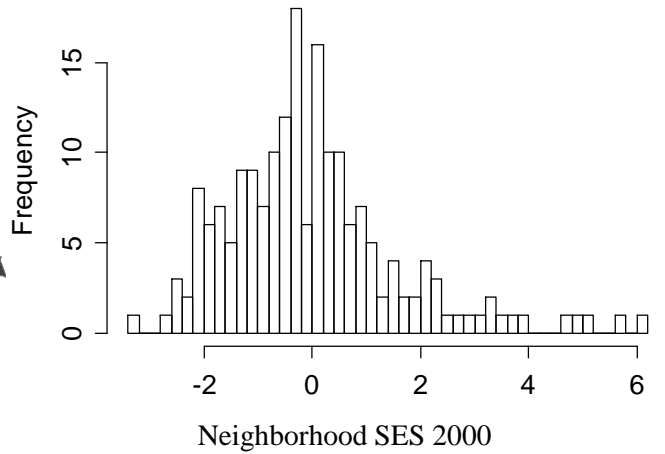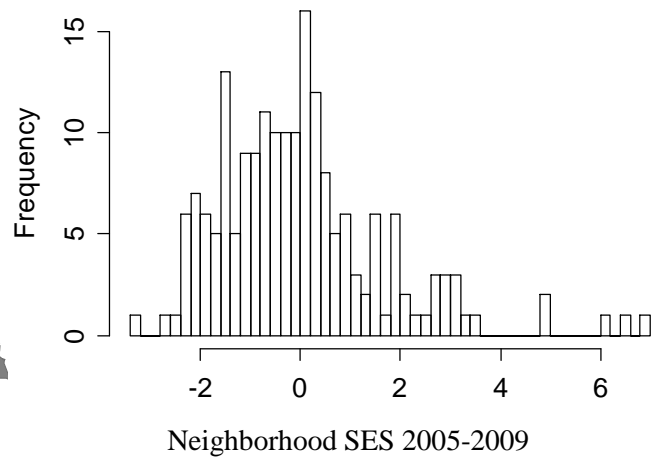
Figure 3.1 Neighborhood SES Maps and Histograms

Maps are in quintiles with lighter colors representing higher SES values

I then calculated the neighborhood percentage white and neighborhood SES for each year by assuming a linear transition from 2000 to 2007 (the halfway point in the ACS 2005-2009 data).

Below are the summary statistics for changes between the census years, 1990 is included for reference purposes. The SES portion of the table show the census-to-census correlations in tract SES. As can be seen, tracts became less white across time and tracts tended to maintain their relative position in terms of SES.

Table 3.1 Changes in (1990) Tracts

	1990 – 2000	2000 – 2005-2009	1990 – 2005-2009
Neighborhood % White			
Mean Change in %White	-6.0%	-3.5%	-9.5%
Median Change in %White	-4.4%	-3.5%	-6.6%
Minimum Change in %White	-39.9%	-24.3%	-46.3%
Maximum Change in %White	26.2%	31.3%	35.3%
Neighborhood SES			
Correlation Coefficient	0.971	0.954	0.911

Because 2003 and 2004 HMDA data had tracts identified in both 1990 and 2000 geographies, they were matched against both census files. Thus, there were some duplicated records. Due to slight changes in census geographies, even within tracts of the same name, the SES variable had slightly different values across duplicates (however, the correlation between these two numbers was over 0.999). To account for this slight discrepancy, I took the average scores when I removed the duplicate tracts.

The following tables list the variables used in the analysis and their corresponding descriptive statistics. They are grouped into categories following the order in which they are introduced into the SHLM.

Table 3.2 Level-1 Descriptive Statistics

VARIABLE	Abbr.	N	MEAN	SD	MIN	MAX
Dependent						
% of Loan Applicants who are White	PCTW	2297	82.32	19.73	0	100
Year						
Year (centered)	YEAR	2297	0.19	3.07	-5	5
YEAR-squared	Y2	2297	9.46	8.78	0	25
Applicant SES						
Natural log of Median Income	LNMINC	2297	4.06	0.35	3.24	5.48
Natural Log of Median Loan	LNMLOAN	2297	4.92	0.34	4.03	6.33
LNMINC-squared	LNMINC2	2297	16.57	2.99	10.48	30.02
LNMLOAN-squared	LNMLOAN2	2297	24.33	3.47	16.24	40.05
YEAR x LNMINC Interaction	YINC	2297	0.74	12.42	-25.57	25.72
Change in Neighborhood Percent White						
Change in Tract % White since 2000	CPW	2297	-3.11	7.11	-34.7	37.18
CPW-squared	CPW2	2297	60.14	141.52	0	1382.58
CPW-cubed	CPW3	2297	-549.27	3998.33	-41780.5	51408.71
Change in Neighborhood SES						
Change in Tract SES since 2000	CNSES	2297	0	0.42	-3.38	3.03
CNSES-squared	CNSES2	2297	0.18	0.62	0	11.39
CNSES-cubed	CNSES3	2297	0.06	1.57	-38.46	27.79

Table 3.3 Level-2 Descriptive Statistics

	VARIABLE	N	MEAN	SD	MINIMUM	MAXIMUM
Neighborhood Racial Composition						
Tract % White in 2000	NW	265	74.97	24.81	0.71	98.91
NW-squared	NW2	265	6234.43	2873.98	0.51	9782.62
NW-cubed	NW3	265	536520.4	299066.2	0.36	967570.1
Neighborhood SES						
Tract SES in 2000	NS	265	0.16	1.68	-2.58	6.19
NS-squared	NS2	265	2.85	5.73	0	38.33
NS-cubed	NS3	265	7.26	30.61	-17.1	237.3
SES Race Interaction						
NW x NS Interaction	NWNS	265	35.27	139.29	-167.68	594.44
Racial Composition Dummies						
Tract 0%-35% White	W35	265	0.09	0.29	0	1
Tract 35%-80% White	W80	265	0.32	0.47	0	1
Tract 90%-100% White	W90	265	0.37	0.48	0	1

Figure 3.2 shows the distribution of the dependent variable: the percentage of applicants in a given tract in a given year that are white. While there is diversity in the proportion of white applicants, the majority of tracts have 80% or more applicants who are white. The relationship between the dependent variable and the key predictor, neighborhood percent white, can be seen in Figure 3.3. There is a strong relationship between these two variables, with a correlation of 0.79. Figure 3.4 presents a scatterplot of neighborhood socioeconomic status and the percentage of applicants who are white. The correlation is strong, but not as strong as neighborhood percent white, at 0.41. The bivariate relationships suggests that racial composition has a stronger relationship with number of applicants who are white than economic status, although Figure 3.4 depicts a non-linear relationship between SES and the percentage of applicants who are white. This

trend is largely mirrored in Figure 3.5 where neighborhood racial composition is plotted

against neighborhood SES. For reference purposes, a histogram of neighborhood percent

white is also included (Figure 3.6).
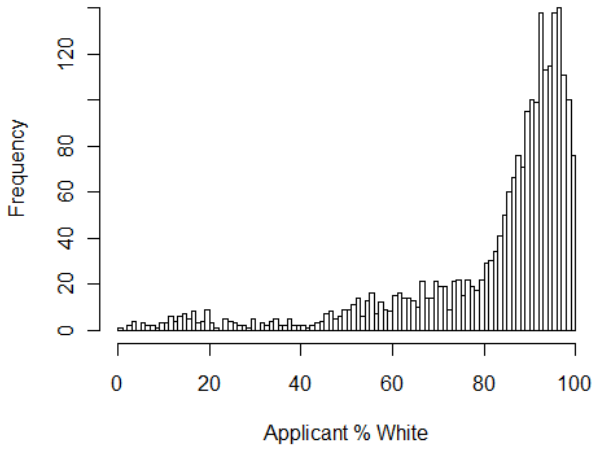
Figure 3.2 Applicant % White
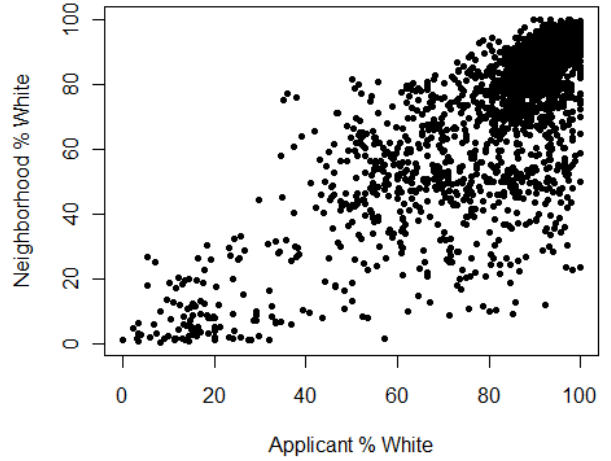
Figure 3.3 Applicant % White by Neighborhood % White
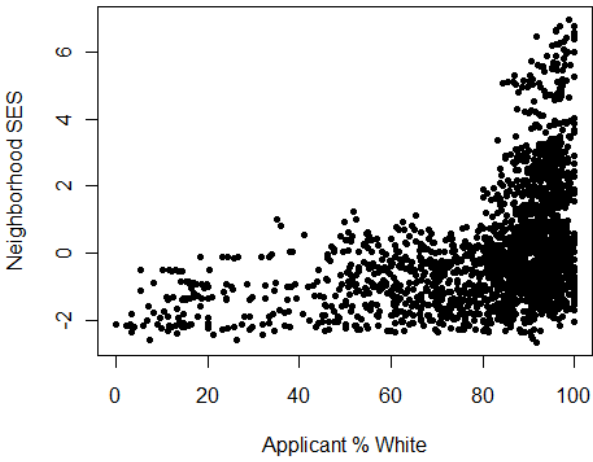
Figure 3.4 Applicant % White by Neighborhood SES
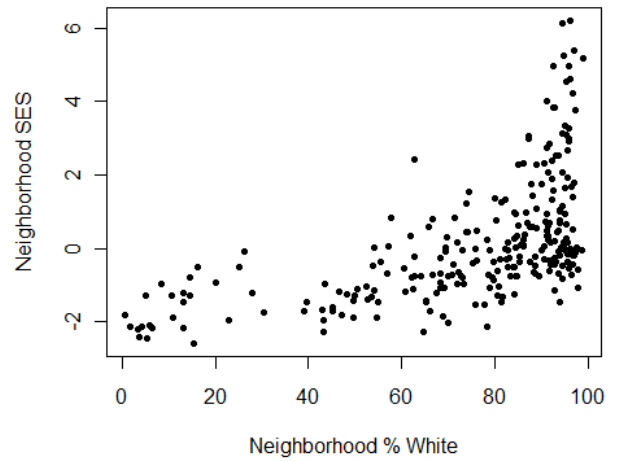
Figure 3.5 Neighborhood % White by Neighborhood SES
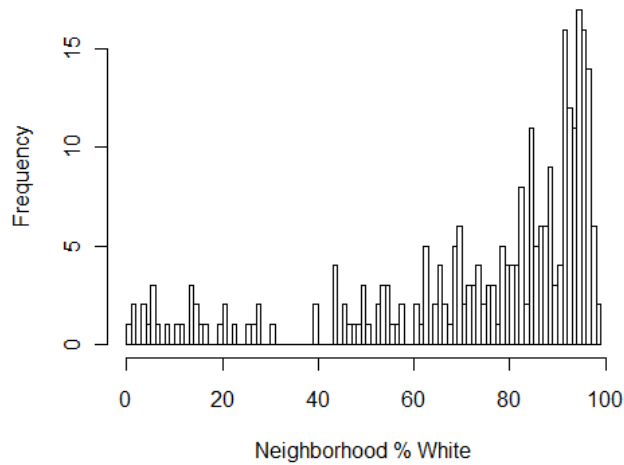
Figure 3.6 Histogram of Neighborhood % White
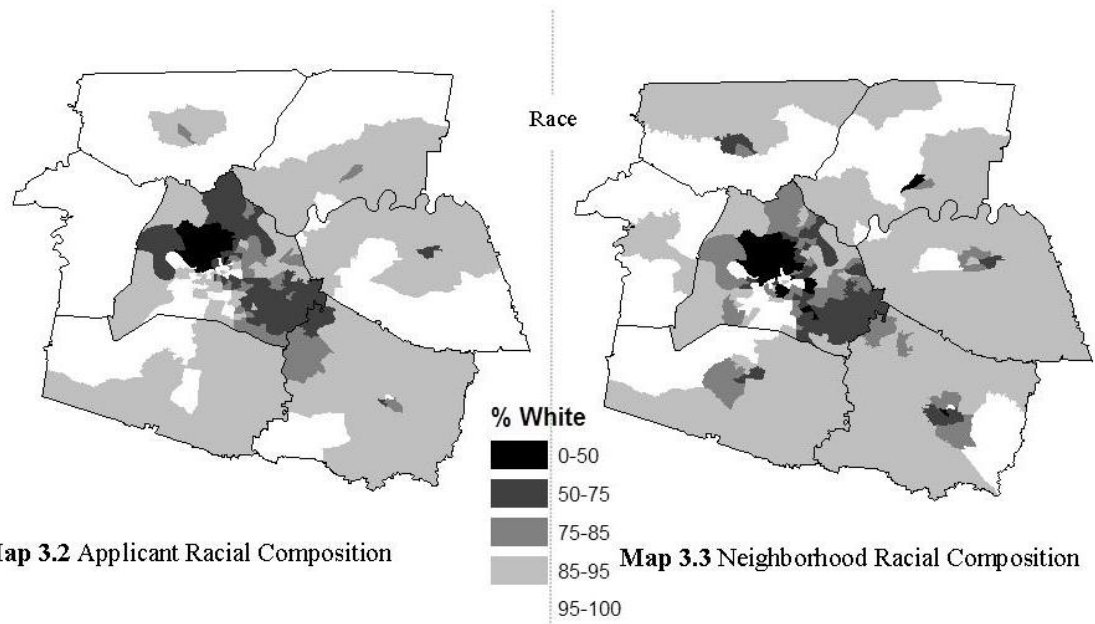
Map 3.2 Applicant Racial Composition

Map 3.3 Neighborhood Racial Composition

% White
0-50
50-75
75-85
85-95
95-100

Race

Applicant

Neighborhood

SES
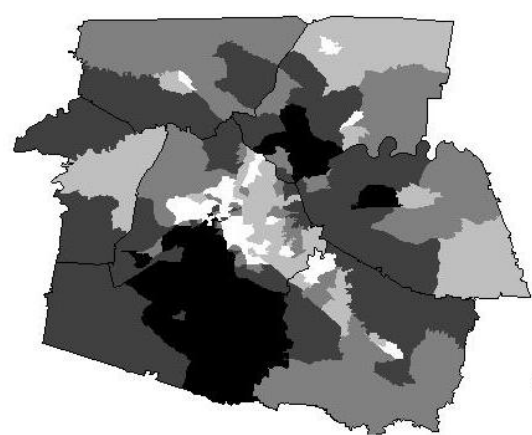
Map 3.4 Applicant Income (in Quintiles)
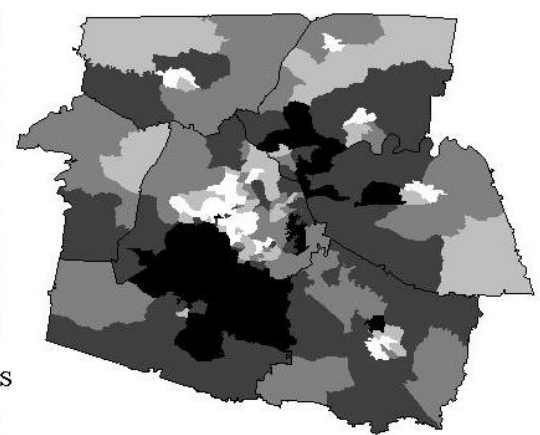
Map 3.5 Neighborhood SES (in Quintiles)

Maps 3.4 and 3.5 are in quintiles with darker colors representing higher Income/SES values

The maps of key variables tell a similar story to that of the scatterplots. The four

maps above (Maps 3.2-3.5) are arranged in quadrants with applicant characteristics on the

left and neighborhood characteristics on the right and racial characteristics at the top and

socioeconomic characteristics at the bottom. One can see that the percentage of

applicants who are white (Map 3.2) follows closely to the neighborhood racial composition (Map 3.3). The socioeconomic status maps (Maps 3.4 and 3.5) also follow the overall pattern of percentage of applicants who are white, although not quite as closely as the racial composition. To look more closely at changes over time, Maps 3.6 through 3.16 show the dependent variable, percent of loan applicants who are white, for each year of the study. These maps seem to suggest that each different region of the city has relatively consistent proportion of applicants who are white.

On the whole, the bivariate relationships explored thus far suggest that there seems to be a stronger relationship between would-be in movers' race and neighborhood racial composition, and a notable, but less strong, relationship between applicant race and neighborhood SES. This suggests that socioeconomic differences are accounting for some, but not all, of the locational patterns of households by race.
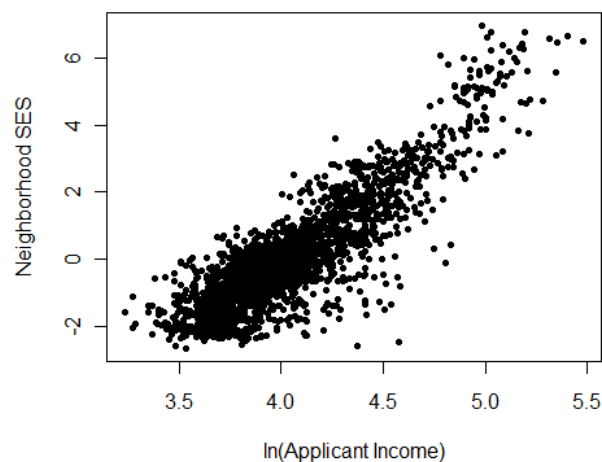
Figure 3.7 Applicant Income by Neighborhood SES

Figure 3.7 shows that the income of applicants was even more strongly related to neighborhood SES than the racial composition of applicants to the neighborhood racial composition (0.87 versus 0.79). This potentially lends some credence to theories that suggest economic similarity is now more important than racial similarities in housing location decisions (c.f., Brown and Chung 2008).

Non-linear Variables

Given that Brown and Chung (2008) have recently argued that the relationship between socioeconomic status is becoming increasingly important, I have included an interaction term between year of application and the median income variable[9]. This interaction term should also help account for potentially changing dynamics in the housing market during the second half of the study period with the onset of the housing crisis (the squared year term should also help account for a non-linearity of the influence of year of application). Similarly, squared terms are included for the applicant socioeconomic status variables, as prior research does not indicate with the impact of these variables would be linear or not.

Since the focus of this study is the connection between neighborhood variables and the percentage of applicants who are white, linear, squared and cubed terms of both neighborhood percent white and neighborhood SES are included in the model. Given the non-linear relationship between neighborhood SES and neighborhood racial composition (illustrated in Figure 3.4), an interaction term of these variables was included. Further, to ensure that the relationship between neighborhood racial composition and the dependent

---

[9] Following convention, the income and loan variables were log transformed to produce a more normal distribution.

variable was modeled adequately, I also created dummy variables delineating specific

racial compositions. The first dummy variable was between zero and 35% white, as

Figure 3.6 suggests that there is a natural break in the distribution of neighborhoods

between 30% and 40% white. I term these neighborhoods "minority neighborhoods"

given their disproportionate share of non-whites. Since the overall racial composition of

the area was approximately 80% white, neighborhoods with between 35% and 80% white

represented another group, which can be described as mixed neighborhoods. Finally,

neighborhoods with more than 90% of the residents white were deemed traditional white

neighborhoods and represent the other dummy variable used in the models

(neighborhoods with 80%-90% white were the reference group).
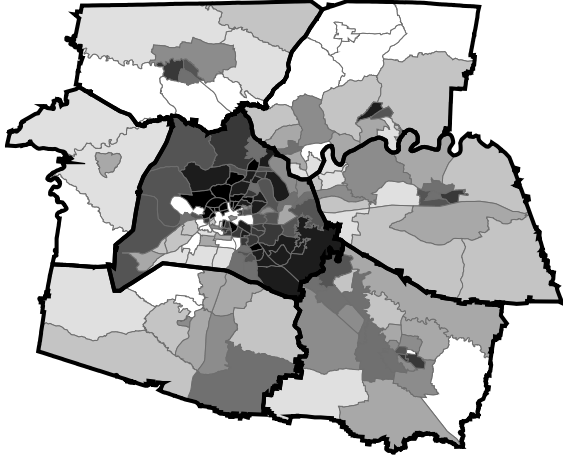
Map 3.6 Applicant % White, in deciles, 2000
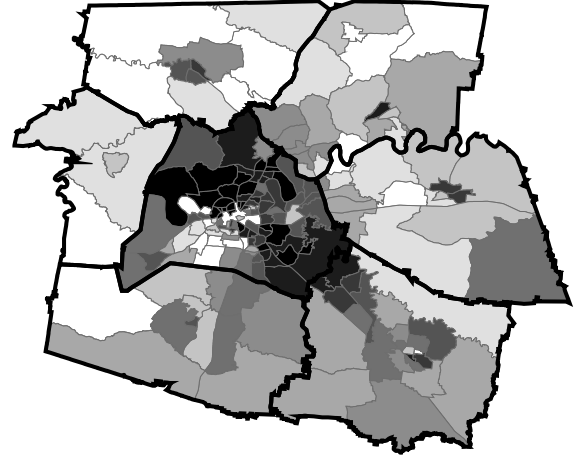
Map 3.9 Applicant % White, in deciles, 2003
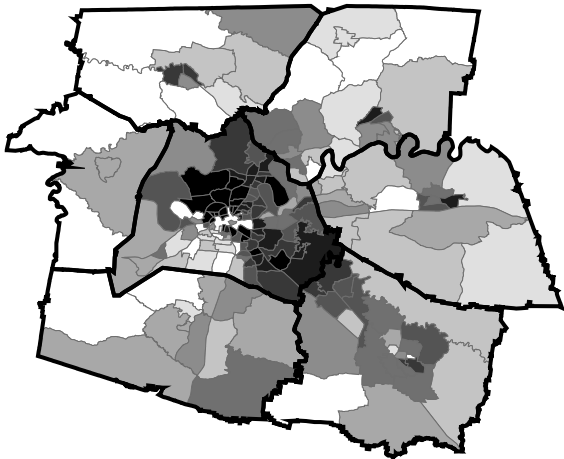
Map 3.7 Applicant % White, in deciles, 2001
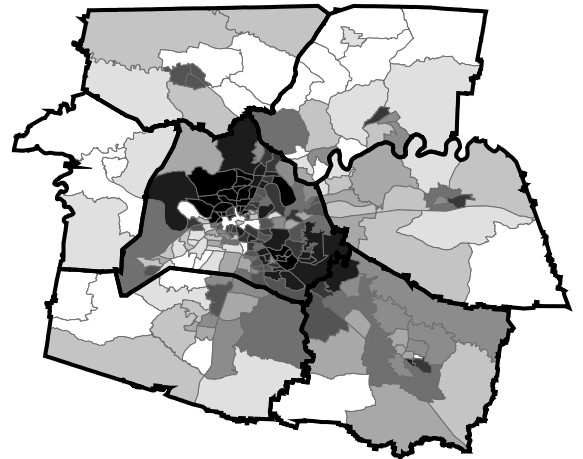
Map 3.10 Applicant % White, in deciles, 2004
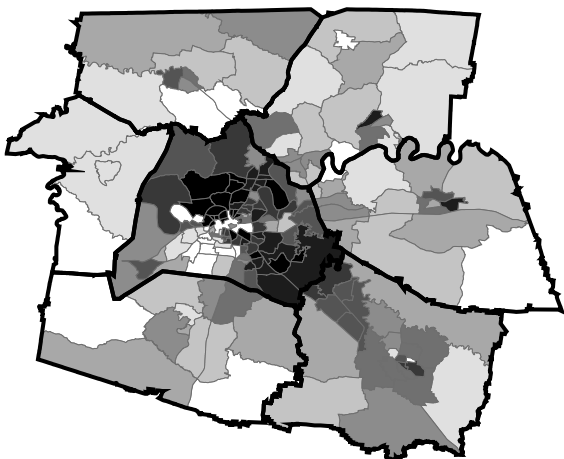
Map 3.8 Applicant % White, in deciles, 2002
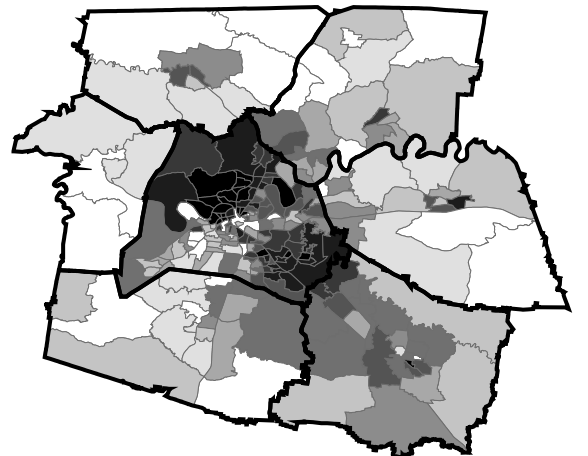
Map 3.11 Applicant % White, in deciles, 2005
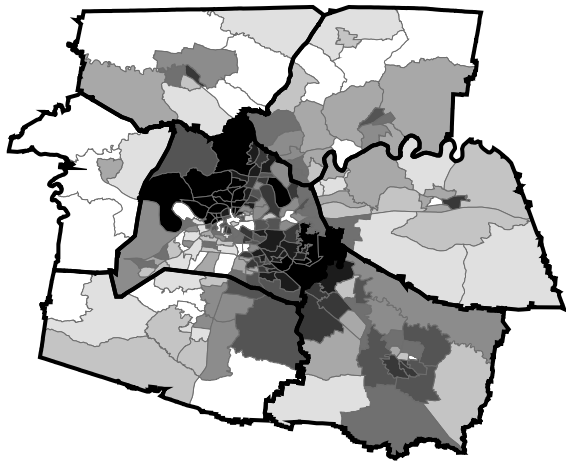
Map 3.12 Applicant % White, in deciles, 2006
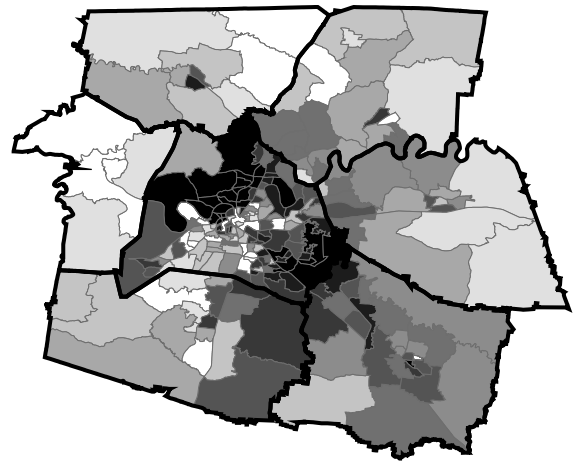
Map 3.15 Applicant % White, in deciles, 2009
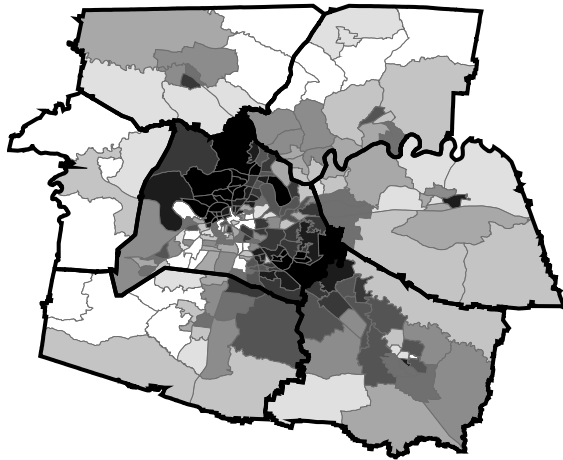
Map 3.13 Applicant % White, in deciles, 2007
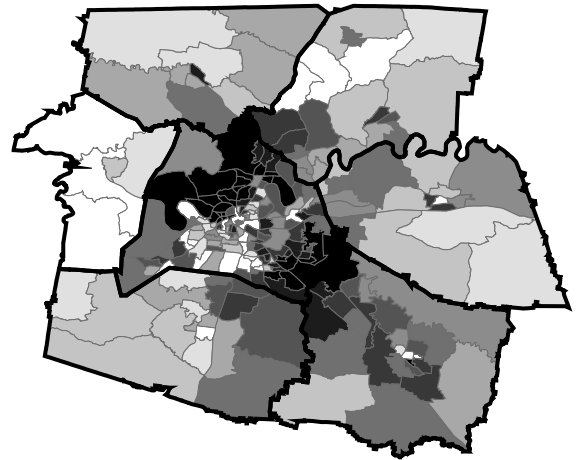
Map 3.16 Applicant % White, in deciles, 2010
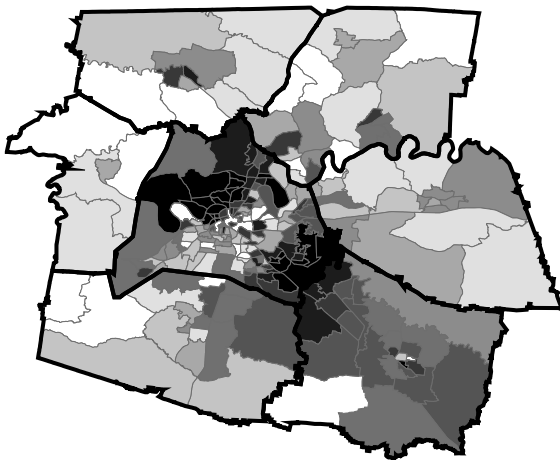
Map 3.14 Applicant % White, in deciles, 2008

Analysis

The analysis used is a spatial hierarchical linear model (Savitz & Raudenbush, 2009). SHLM provides an elegant analytical tool for this research as it incorporates the information provided by repeat observations within each tract as well as the spatial relationships[10] between tracts.

Analysis Plan

In their Hierarchical Linear Models text, Raudenbush and Bryk (2002) highlight how developing a model is both an art and a science. They write: "The early phases of model building involve an interplay of theoretical and empirical considerations. The substantive theory under study should suggest a relatively small number of predictors for possible consideration in the level 1 model" (p. 256). Following Raudenbush and Bryk (2002), the analysis begins by looking at the unconditional model (with no predictors) and add the substantive predictors in groups in order to develop the most parsimonious predictive model of the percentage of applicants who are white. Theoretically linked variables will be added to the model in groups[11]. Each variable within the group will be assessed as to whether it helps improve the model. Given the complexity of multilevel modeling, the overall goal is to create a theoretically derived parsimonious model. Thus, variables with little theoretical and substantive importance and low predictive power ($p > 0.15$) will be

---

[10] Specifically, the SHLM accounts for the correlation between error terms of adjacent tracts.
[11] This is based on the suggestion of Prof. Tom Smith.

dropped in the model building process[12]. In this study, groups of variables will be added

in the following order (this is described in more detail below):

1. Unconditional model (no predictors)

2. Unconditional growth model (year variables)

3. Socioeconomic characteristics of applicants (income and loan amounts)

4. Neighborhood racial composition variables

5. Socioeconomic characteristics of neighborhoods

6. Changes in neighborhood racial composition

7. Changes in neighborhood socioeconomic status

8. Neighborhood racial composition dummy variables

The unconditional model provides variance estimates that allow for the importance of

time-varying versus stable characteristics to be assessed (this is known as the intraclass

correlation [ICC]). The next step – controlling for the impact of time – is the next stage

in any growth model. Raudenbush and Bryk (2004) assert that the level-1 predictors

should be added to the model first, and so the loan and income variables are introduced

next (first just the linear and then the squared terms) and then the interaction term

between year and income. When the best model has been specified at level-1, including

year and socioeconomic covariates, then the racial composition variables at level-2 are

included. This is followed by the socioeconomic neighborhood variables. It is then

assessed whether there is an interaction between socioeconomic status and neighborhood

racial composition. After the most parsimonious model has been developed, the racial

---

[12] While it is clear that certain constructs may be important predictors of the percentage of applicants who are white (e.g., median income or neighborhood SES), it is difficult to predict the exact form and fit of this relationship (hence the non-linear and interaction terms included in the models).

change variables at level-1 are added (these would have been relatively meaningless without the level-2 variables included first). After assessing the importance of these variables, the change in neighborhood socioeconomic status variables are included. Finally, as a further check to ensure that the relationship between neighborhood racial composition and the percentage of applicants who are white is being adequately modeled, dummy variables denoting key points in the distribution of neighborhood racial composition are added. In its simplest form[13], the final model would look like this:

Level-1 equation:

$$PCTW = \pi_{0n} + \pi_{1n} \, YEAR + \pi_{2n} \, LNINC + \pi_{3n} \, LNLOAN +$$
$$\pi_{4n} \, CPW + \pi_{5n} \, CNSES + e_{jn}$$

Level-2 equations:

$$\pi_{0n} = \beta_{00} + \beta_{01} \, NW + \beta_{02} \, NS + r_0$$
$$\pi_{1n} = \beta_{10}$$
$$\pi_{2n} = \beta_{20}$$
$$\pi_{3n} = \beta_{30}$$
$$\pi_{4n} = \beta_{40}$$
$$\pi_{5n} = \beta_{50}$$

Spatial dependence equation:

$$r_0 = pW \, r_0 + b_0$$

The level-1 equation depicts how the dependent variable (percentage of applicants who are white for a given tract in a given year) is to be predicted by an intercept value

---

[13] Without squared or interaction terms.

(π_{0n}) and the year of application, the median income, the median loan amount, the change

in neighborhood percent white since 2000, and the change in neighborhood SES since

2000. The impact of each of these independent variables is accounted for in the parameter

estimates of π_{1n} through π_{5n}. As can be seen from the level-2 equations, the parameter

estimates for π_{1n} through π_{5n} will be the same for each tract. The intercept (π_{0n}),

however, will be estimated based on the neighborhood racial composition (NW) and

neighborhood SES (NS) in the year 2000. This means that the predicted value for the

intercept will potentially change for each neighborhood in the sample. The spatial

dependence equations shows how the level-2 error on the intercept (r_0) will be

decomposed into two components, neighborhood level random error (b_0) and spatially

dependent aspect of the data. In this equation, p is the spatial correlation, W is the spatial

weight matrix, and r_0 is a vector of random spatially autoregressive effects.

Results

Model 1: Unconditional Model

The unconditional model (no independent variables) allows for the calculation of

the intraclass correlation (ICC). The ICC estimates the amount of variation in the

dependent variable associated with level-2. The ICC from the non-spatial HLM

attributes 87% of the variation in the percentage of applicants who are white to between

tract differences, with the remaining 13% associated with within tract changes. However,

when spatial dependence is modeled through the SHLM, the tract level predictors are reduced to 58% of the variance, as the spatial correlation of adjacent tracts in the dependent variable is a striking 0.95.

Thus, we can conclude, as would have been expected from examining Maps 3.6 through 3.16, a combination of spatial and tract level factors are far more important in predicting the percentage of applicants who are white than are any factors within a tract that change over the study period. This suggests that, when it comes to racial composition of loan applicants, the 11-year study period is not long enough for major changes to occur in most census tracts in the Nashville area.

Model 2: Looking at the importance of time

Both time variables (YEAR and Y2) were positive and significant, suggesting that the proportion of applicants who were white increased across the study period in all tracts. Including the time level variables accounted for 14% of the variance associated with level-1. As can be seen in Figure 3.8, there was a predicted 9% difference between proportion of applicants who were white in 2000 versus 2010.
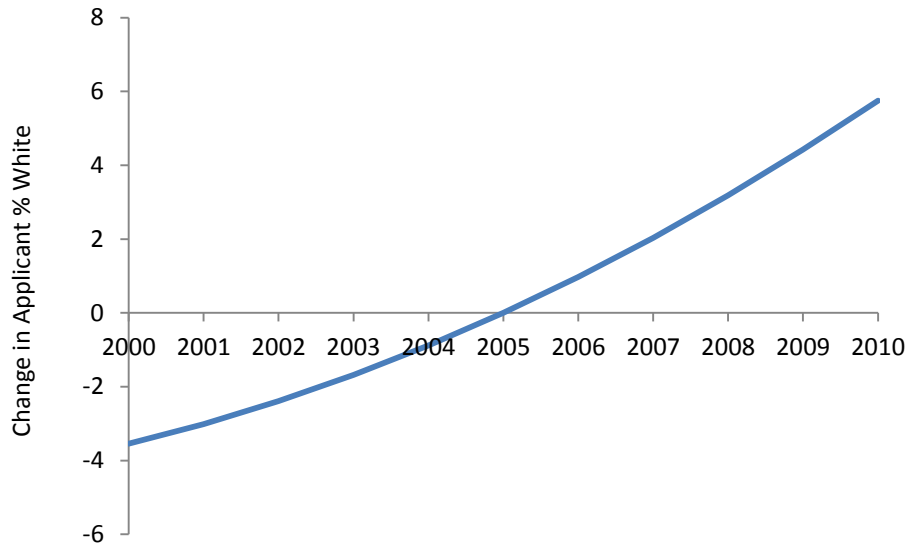
Figure 3.8 Impact of Year Variables on % of Applicants who are White

Model 3: Introducing level-1 socioeconomic predictors

The income and loan variables positively predicted the percentage of applicants who were white (after controlling for time). The cumulative impact of a one standard deviation change in these variables is displayed in Figure 3.9. This figure shows that a one standard deviation increase in each variable increases the percentage of applicants who are white by about 1.7%. So a neighborhood that had a median loan amount and median income of one standard deviation higher than the average neighborhood would be predicted to have 3.4% increase in applicants who were white.
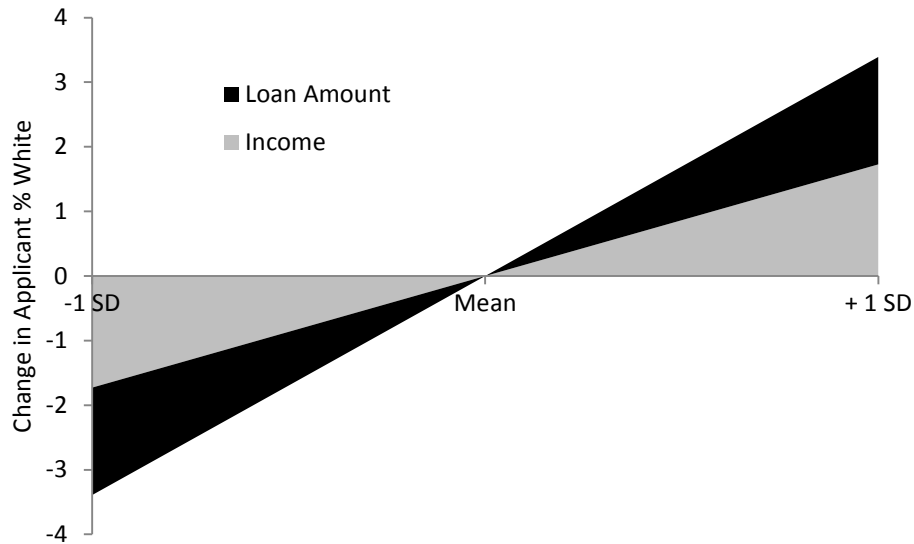
Figure 3.9 Cumulative Impact of Income and Loan amount Variables on % of Applicants who are White

The natural log of median loan squared (LNMLOAN2) was also significant, but the natural log of income-squared (LNMINC2) was not. Both the non-squared loan and income variables were positive and significant, with loan-squared significant and negative. This fuller model, including the two loan variables and the one income variable, explained 19% of level-1 variance, and 9% of the level-2 variance (this suggests that the level-1 socioeconomic aspects are relatively consistent across time and including these variables reduces the amount of variation associated with level-2). An interaction term between the median income of applicants and the year of application was introduced to assess whether the impact of income changed across time. This variable was strongly significant and negative, suggesting that the positive relationship between applicants' income and the percentage of applicants who were white declined across time. This provides some evidence that socioeconomic factors are of declining importance when

explaining racial residential segregation. Including this interaction term increased the total level-1 variance explained to 24%.

Model 4: Introducing neighborhood racial composition

Next, the association between neighborhood racial composition and the percentage of applicants who were white was tested (after controlling for year and the level-1 median income variables). All three neighborhood racial composition variables were strongly statistically significant, with both NPW and NPW3 having positive coefficients and the squared term (NPW2) having a negative coefficient. Figure 3.10 shows the impact of the three neighborhood racial composition variables on the percentage of applicants who are white. The gray dotted line demarks a percentage of applicants who are white equal to the racial composition. The graph clearly shows the opposite of a Schelling-type tipping. Instead, it shows a strong tendency towards integration. A disproportionately greater (in relation to a 1:1 relationship of NPW and PCTW: the dotted line) number of white households are drawn to minority neighborhoods. According to the results of this model, minority neighborhoods would get whiter, and white neighborhoods become less white. Including these variables explained 44% of the total level-2 variation and reduced the spatial correlation coefficient by 0.07 to 0.89. Table 3.4 shows the full results of this model.

.
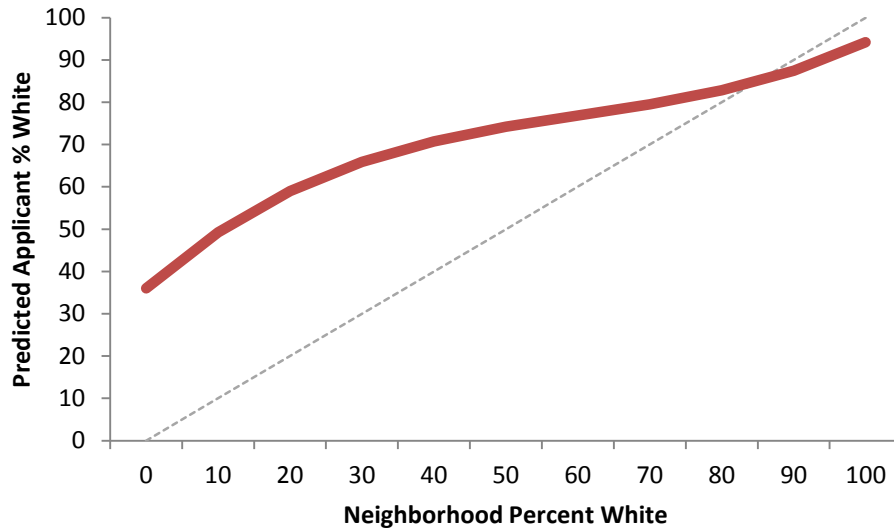
Figure 3.10 Relationship between Neighborhood % White and Applicant Racial Composition (Model 4)

Table 3.4 Final Results for Model 4, 5 and the Final Model

Fixed Effect	Model 4		Model 5		Final Model	
	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
INTRCPT	35.98607	<0.001	23.30635	<0.001	-	-
NW	1.520703	<0.001	1.719387	<0.001	1.972785	<0.001
NS	-	-	-7.59945	<0.001	-12.1325	<0.001
NW2	-0.02089	<0.001	-0.02248	<0.001	-0.01805	<0.001
NW3	0.000115	<0.001	0.000128	<0.001	0.00008	0.008
W35	-	-	-	-	12.84962	<0.001
NWNS	-	-	0.042226	0.085	0.098477	<0.001
YEAR	8.580494	<0.001	9.949781	<0.001	10.70397	<0.001
Y2	0.076433	<0.001	0.074812	<0.001	0.047862	0.005
LNMINC	8.786581	<0.001	9.814649	<0.001	9.064786	<0.001
LNMLOAN	53.79351	0.002	4.527549	0.004	-	-
LNMLOAN2	-5.3045	0.003	-	-	-	-
YINC	-1.91236	<0.001	-2.27178	<0.001	-2.44248	<0.001
CPW	-	-	-	-	0.252534	<0.001
CNSES	-	-	-	-	2.330402	<0.001
CPW2	-	-	-	-	0.007387	<0.001
CNSES2	-	-	-	-	1.52836	<0.001

Model 5: Neighborhood SES

      Since neighborhood racial composition and SES are correlated, the next step introduced the neighborhood socioeconomic status as control variables. Only the linear term was statistically significant. Perhaps surprisingly, it was also negatively related to the percentage of applicants who were white. The interaction term between neighborhood racial composition and neighborhood SES (NWNS) was added next and it was statistically significant and positive. This suggests that the strength of the relationship between neighborhood percent white and the percentage of applicants who are white becomes stronger as neighborhood SES increases. At level-1, loan-squared became a non-significant predictor when the socioeconomic status variables were included at level-2, so this was dropped from the model. The effect of racial composition was slightly less pronounced than it was in Model 4 after NS and NWNS were included. Figure 3.11 demonstrates this for when NS (and therefore, NWNS) is set to zero. This model accounts for 53% of the level-2 variation, with the level-1 variance explained remaining at 24% and the spatial correlation substantially unchanged at 0.885. Table 3.4 provides the parameter estimates for the model.
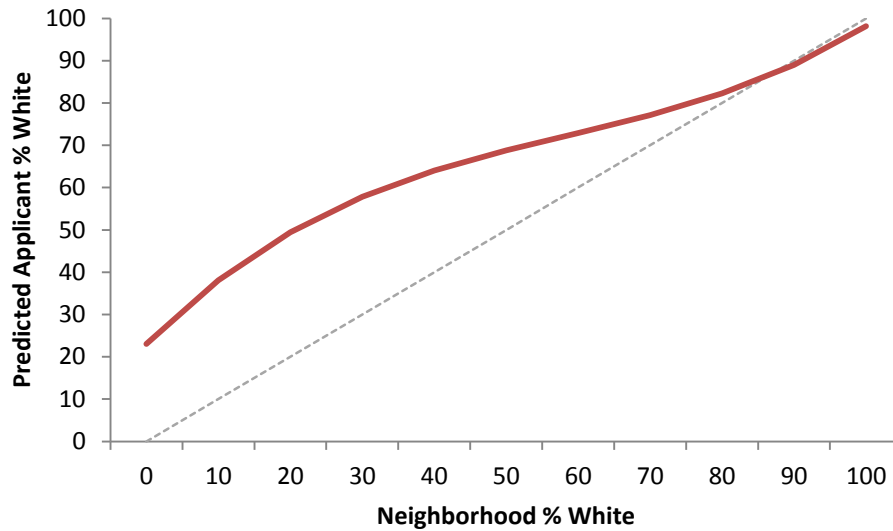
Figure 3.11 Relationship between Neighborhood % White and Predicted Applicant % White controlling for neighborhood SES (Model 5)

Models 6 and 7: Neighborhood Change

The neighborhood racial composition variable included until now was for the year 2000. Changes in neighborhood racial composition (CPW) may be an important aspect of the decisions of would be movers. When the CPW variables were included, CPW3 was not significantly predictive of the outcome variable, but both CPW and CPW2 were significant and positively associated with percentage of applicants who were white. The effect of this variable was moderate. For example, a 5% increase in tract percentage white would result in a predicted increase of the percentage of applicants who were white of 2%. A 20% change in neighborhood racial composition would result in an increase of the percentage of applicants who were white by approximately 11%. This model improves upon earlier models and accounts for 25% of level-1 variation and 62.5% of level-2 variation.

The change in neighborhood SES variables were included next. As with the change in neighborhood racial composition variables, the cubed term for change in neighborhood SES (CNSES3) was not related to the outcome variable, but the linear and squared terms were. Loan amount ceased to be a significant predictor and was dropped from the model. This model increased the predictive power of the level-1 variation to 26.4% but reduced the level-2 variation slightly to 61%.

Model 8: Minority, Mixed, and All-White Neighborhoods

Finally, I introduced the three dummy variables representing four groupings of neighborhood racial composition. As mentioned earlier, the first grouping was what could be termed "minority neighborhoods", neighborhoods that are less than 35% white. There exists few neighborhoods in the 30-40% white range (see Figure 3.6), so this seemed to be a meaningful gap in the distribution of neighborhoods (just under 10% of the neighborhoods in the sample met this criteria). The second grouping was 35-80% white, which could be termed "mixed neighborhoods": any neighborhood that has less than a proportional representation of whites but greater than the 35% cut-off (almost a third of neighborhoods met this criteria). The third (and reference group) was those neighborhoods that had between 80 and 90% white (approximately 22% of all neighborhoods). The final grouping was "white neighborhoods" that had tract percentage white over 90% (37% of the neighborhoods met this criteria).

When these 3-dummies were added to the model, only the minority neighborhood dummy variable was significant (it was positively associated with the percentage of applicants who were white). This final model accounted for 26.3% of level-1 variation,

62.6% of level-2 variation and had a spatial correlation in the dependent variable of 0.9. The results of this model is shown in the last column of Table 3.4.

Figure 3.12 illustrates the results of the final model. It shows the predicted percentage of applicants who are white (y-axis) based on the percentage of the tract population that is white (x-axis). The light gray line is the predicted value based on just the level-2 racial composition variables (everything else set to zero). The darker gray line is the predicted value when all level 2 variables are included in the model at their means for that particular racial composition. The black line represents the predicted value when all the level one predictors are also included at their mean for that racial composition. This graph attempts to show a realistic depiction of racial composition of in-movers based on the typical profile of applicants moving into a neighborhood of that racial composition.

All lines show higher-levels of white homeowners than the overall racial composition when the neighborhood percent white is below 80. Between 80 and 90 percent white the predicted percentage of applicants who are white becomes approximately equal and then at the highest percentage of white residents, it dips slightly below. While the relationship is not as clear as it was in Figure 3.6, it still demonstrates that white loan applicants are not following a Schelling-type pattern in their neighborhood selection.
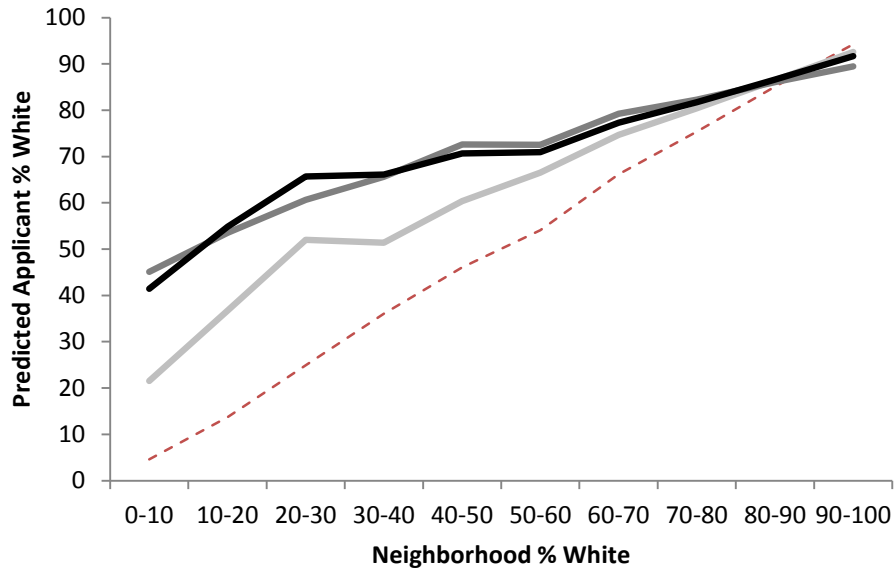
Figure 3.12 Relationship between Neighborhood % White and Applicant Racial Composition, Controlling for Neighborhood and Applicant Factors (Model 8)

Discussion

The results of Fossett's and Schelling's theoretical models provide a compelling case for preference based models of segregation. Two key empirical findings in this chapter provide further support for the preference model of segregation. First, as neighborhood percentage white increases, there is a general tendency for applicant percentage white to increase. As shown in Figure 3.12, in the least white neighborhoods, 40% of the applicants are white, while in the whitest neighborhoods, 90% of the applicants are white. Second, both neighborhood racial change variables were positive and significant. This suggests that applicants are responsive to neighborhood racial change, and white applicants are more likely to apply to live in neighborhoods that are becoming whiter.

However, when the full results of the analysis are looked at, it seems clear that there is little evidence for neighborhood tipping among would-be homeowners. White homebuyers make up at least 50% of the applicants in nine out of ten of the neighborhood racial composition groupings depicted in Figure 3.8. Further, the only neighborhoods that have approximately equal or lower proportion of white applicants than white residents are the neighborhoods that are between 80 and 100 percent white. Fossett's and Schelling's model predict the opposite of this.

If the preference models of segregation do not provide a satisfactory explanation for the results presented above, can socioeconomic factors explain the patterns revealed in the data? The short answer is, probably not. A one standard deviation increase in applicant income predicts an increase in applicant percentage white of 3% while a one standard deviation increase in neighborhood SES actually predicts a 20% decrease in applicant percent white. This suggests there is a moderate relationship between applicant income and the percentage of white applicants moving into a tract, but at the same time a strong negative relationship between neighborhood SES and the proportion of whites moving into that neighborhood. The positive coefficient on the neighborhood SES x Race interaction term counter balances the negative impact of neighborhood SES to some extent. Higher SES neighborhoods are associated with a stronger positive relationship between the neighborhood percent white and the percentage of applicants who are white. That is, neighborhood SES accentuates the impact of neighborhood racial composition. The positive impact of SES only occurs through the interaction with neighborhood racial composition. If SES factors mediated the relationship between applicant race and destination neighborhood, we would expect neither a negative coefficient on

neighborhood SES nor a significant interaction between neighborhood SES and neighborhood racial composition.

This study cannot test for the importance of racial biases of housing institutions, but the lack of compelling evidence for the two other major explanations of segregation suggests that it would be premature to dismiss discrimination-based theories of racial residential segregation at this particular point in time.

This study highlights the empirical weakness of preference and socioeconomic explanations for contemporary residential racial segregation among households looking to purchase a home in the Nashville region over the past decade. This study does not examine patterns in the rental market and it could be that preference or socioeconomic theories provide empirically compelling explanations of locational decisions of renters. Nor does the study examine the racial composition of households leaving neighborhoods, it could be that the pattern of leavers is as, or more, determinative of neighborhood racial composition as the racial composition of in-movers. Both of these aspects of the housing market, movement patterns of renters and racial composition of households leaving neighborhoods, would be extremely productive areas for future research given the results described in this study.

CHAPTER IV

DO LENDING DECISIONS SHAPE NEIGHBORHOOD RACIAL COMPOSITION?

Chapter 4 analyzes the impact of lending decisions on neighborhood racial composition. Chapter 3 investigated the relationship between the racial composition of applicants and the neighborhoods they attempt to move into, and thus aimed to provide a picture of the relationship between individual moving decisions and racial segregation. This chapter asks a similar question, but instead of focusing on the characteristics of applicants, it focuses on the decisions of lenders and how they relate to neighborhood racial composition. In order to accomplish this, the analysis will assess the extent that loan approval decisions reshape the racial make-up of new residents in each neighborhood by comparing the racial make-up of all loan applicants to the racial make-up of those who are approved by lenders. It asks the question: do the decisions of mortgage lenders alter the racial composition of a neighborhood from what it would be if all applicants (or at least a racially proportionate number) were able to receive financing for their home? Put another way, can current levels of segregation of homebuyers be explained by where households apply to live, or do lending decisions play a key role in shaping the racial landscape of Nashville's metropolitan area? While this analysis may not indicate discrimination, it may show an important factor shaping segregation that preference-based models do not account for. There have been a large number of analyses on discrimination in lending (to be reviewed in Chapter 5), there are, however, no

examinations of the role of lending institutions and their direct impact on residential racial segregation.

The analysis will show the impact of lending institutions on neighborhood racial composition and whether it has changed over time. Specifically, the null hypothesis can be stated:

There will be no significant relationship between tract percentage white and the extent that whites are favored in lending approvals. This will be true even after controlling for neighborhood socioeconomic status and applicant socioeconomic characteristics.

This analysis will use a two level spatial hierarchical growth model to predict the difference in racial composition of successful applicants versus all applicants for each census tract in each year. That is, the dependent variable will be calculated by looking at the percentage of white applicants and then calculating the difference between that and the percentage of successful applicants who were white. For example, let's say 50 white households and 50 non-white households apply to move into census tract A in year X. Of those, 30 whites and 20 non-whites are accepted. Fifty percent of applicants were white (50/100), but 60% of successful applicants are white (30/50), leading to difference of positive 10%[14]. Thus, in this hypothetical neighborhood (tract A), the impact of lending decisions led to the residents being whiter than what would have been if location was determined strictly by where households apply for loans.

---

[14] More formally: Dependent Variable (Percentage of Successful White Applicants in Tract A in Year X – Percentage of Total Applicants who are White in Tract A in Year X) = 60% - 50% = 10%

Similar to Chapter 3, a spatial growth model will be used as the data have a natural nested structure, with each census tract having up to 11 observations (one for each year). That is, we have a measurement of how much white applicants were favored/disfavored for every year for each tract. The data have a spatial component, with some census tracts closer to some and further away from others. Ignoring this spatial factor would potentially lead to misspecified models and downwardly biased standard errors.

The key independent variable of interest will be neighborhood racial composition. To control for socioeconomic factors that may be correlated with race, I will include a variable for median income of white applicants and median income of minority applicants. I will also include median loan amounts in a similar manner. To further control for these socioeconomic factors, the ratio of the incomes (median white income/median minority income) and loan amounts (median white loan/median minority loan) as well as the loan to income ratios of both groups (discussed in more detail below) will also be included. The results of this study will tell us whether lending decisions are playing an important role in segregative patterns (increasing, decreasing, or having no impact on segregation).

Data

The data sources are the same as for Analysis 1 (more detail on HMDA and Census data can be found in Chapter 3). After the loan application data were pulled into R statistical software (n=439,644), I performed the following operations:

1. I removed records that did not have incomes between $1 and $1,000,000. I then adjusted the income and loan amount variables by inflation using the Bureau of Labor Statistics inflation calculator. All years were normalized to 2009 dollar amounts. 425,585 applications remain.

2. I included only cases where the loan was either approved or denied (rather than withdrawn). 326,470 applications remain.

3. I calculated, for each year, and separately for whites and minorities, census tract median loan amount and the median income for the 2,601 year-tracts in the dataset.

4. Since lending decisions are influenced by the borrowers' ability to repay the loan, I created 5 ratio variables to attempt to control for this aspect of the loan approval process.

    a. First, I created a ratio of white income to minority income (INCRAT) and a ratio of white loan amount to minority loan amount (RATIO).

    b. Then I created a variable of the ratio of white loan amount to white income (ILW), and a minority loan amount to minority income (ILM).

    c. The fifth variable was the ratio of these last two ratios (ILR).

5. In order to create the dependent variable, I calculated the total number of applicants and the total number of successful applicants for each census tract in each year. Similarly, I calculated the total number of applicants who were white as well as the number of successful applicants who were white. I then calculated the percentage of successful white applicants in each tract in each year then subtracted the percentage of total applicants who were white for each respective tract and year. This created the dependent variable.

6. I removed any cases that had either no white or no minority applicants or had less than 20 applicants. 2,071 year-tract observations remain.

7. The dependent variable was examined and outliers (those with a more than 10% change in a given year-tract) were removed (28 cases). 2,043 observations remain.

8. I then examined the independent variables and removed any cases that had significant outliers on any of the predictors. A number of the economic characteristics were significant outliers when histograms were examined. Removing these cases facilitates developing a predictive model that has a better fit for most cases, but at the expense of potentially excluding informative outliers. A total of 48 observations were removed (2.3% of the sample) from the loan, income, and ratio variables. 1,995 cases remain.

Table 4.1 shows the descriptive statistics for each of the variables used in the model.

Table 4.1 Descriptive Statistics

Level-1 Variables

	Label	N	MEAN	SD	MIN	MAX
Dependent						
Difference between the % of applicants who were white and the % of successful applicants who were white.	Δ%W	1995	0.9	2.32	-9.13	9.87
Time						
Year centered around 2005	TIME	1995	0.16	3.06	-5	5
TIME-squared	T2	1995	9.41	8.82	0	25
Applicant Characteristics						
Natural log of median white income	MEDWINC	1995	4.06	0.34	3.22	5.25
Natural log of median white loan amount	MEDWLOAN	1995	4.9	0.34	4.02	6.03
Natural log of median minority income	MMINC	1995	4.03	0.44	3.04	5.94
Natural log of median minority loan amount	MMLOAN	1995	4.92	0.41	3.74	6.45
MEDWINC/ MMINC	INCRAT	1995	1.01	0.07	0.73	1.32
MEDWLOAN/ MMLOAN	RATIO	1995	1	0.05	0.85	1.23
MEDWLOAN/ MEDWINC	ILW	1995	1.21	0.05	1	1.47
MMLOAN/ MMINC	ILM	1995	1.23	0.08	0.84	1.57
ILW/ILR	ILR	1995	0.99	0.07	0.8	1.5
Change in Neighborhood Racial Composition						
Change from 2000 neighborhood percent white	CPW	1995	-3.26	7.11	-34.7	44.78
CPW-squared	CPW2	1995	61.18	148.75	0	2005.13
CPW-cubed	CPW3	1995	-561.43	4514.9	-41780.5	89786.92
Change in Neighborhood SES						
Change from 2000 NS	CNSES	1995	-0.04	0.41	-3.03	3.38
CNSES-squared	CNSES2	1995	0.17	0.62	0	11.39
CNSES-cubed	CNSES3	1995	-0.05	1.62	-27.79	38.46

Level-2 Variables

		N	MEAN	SD	MIN	MAX
Neighborhood Racial Composition						
Neighborhood Percent White in 2000	NW	262	75.29	24.45	0.71	98.91
NW-squared	NW2	262	6264.68	2846.87	0.51	9782.62
Neighborhood SES						
Neighborhood Socioeconomic Status in 2000	NS	262	0.15	1.64	-2.58	6.19
NS-squared	NS2	262	2.7	5.35	0	38.33
NS-cubed	NS3	262	6.54	27.53	-17.1	237.3
SES Race Interaction						
NW x NS Interaction	NWNS	262	33.78	135.82	-167.68	594.44
Racial Composition Dummy Variables						
Tract 0%-35% White	W35	262	0.09	0.29	0	1
Tract 35%-80% White	W80	262	0.32	0.47	0	1
Tract 90%-100% White	W90	262	0.37	0.48	0	1

Below is a histogram of the dependent variable. As can be seen, it peaks around zero and is slightly heavier on the positive side than on the negative side (the average of 0.9 suggests that lenders have a tendency to favor white applicants). Below the histogram is a scatterplot of neighborhood percentage white on the x-axis and the dependent variable on the y-axis (figure 4.2). There does not seem to be a clear relationship between these two variables (in fact, the correlation coefficient is -0.16). The relationship between neighborhood socioeconomic status and the dependent variable is also depicted below. There seems to be a very slight favoring of whites in high economic status neighborhoods, and a large amount of variation in low-SES neighborhoods (correlation coefficient is -0.12).
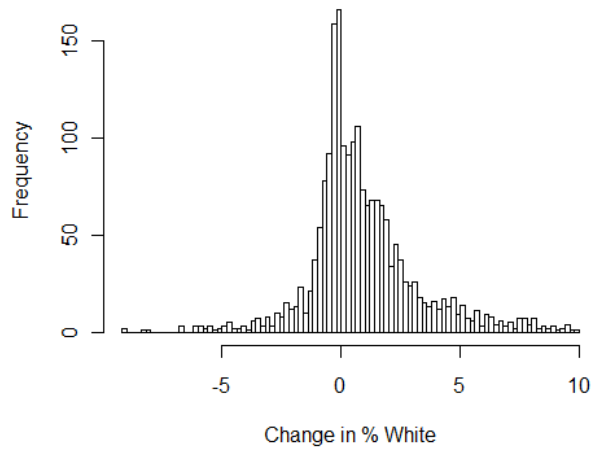
Figure 4.1 Dependent Variable (Δ%W)
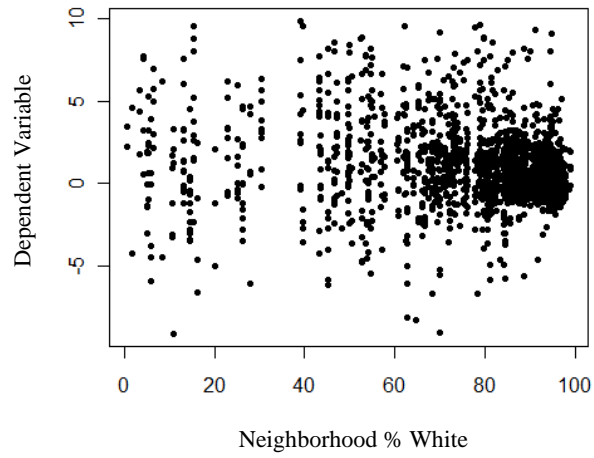
Figure 4.2 Neighborhood % White by Δ%W
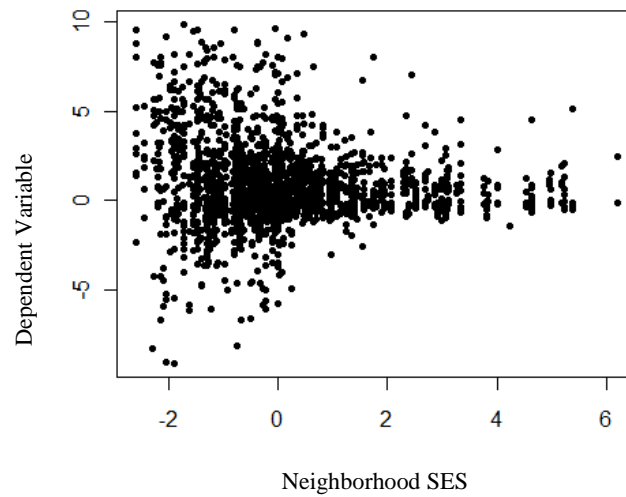
Figure 4.3 Neighborhood SES by Δ%W

The center of Map 4.1 suggests that there might be relationship between

neighborhood racial composition and the decisions of lenders. The area that showed a
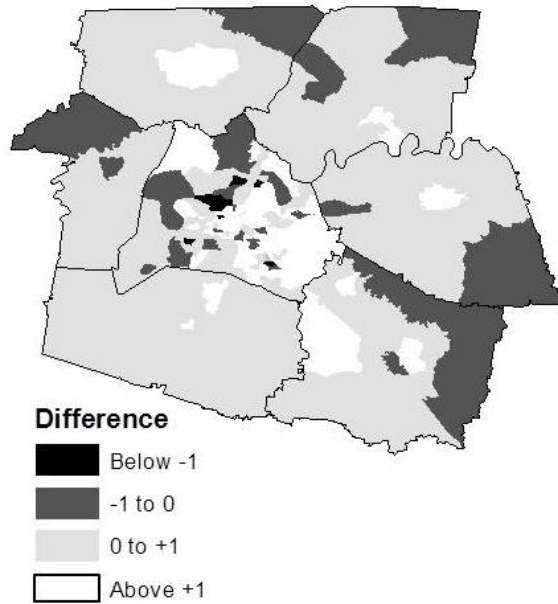
Map 4.1 Average Δ%W
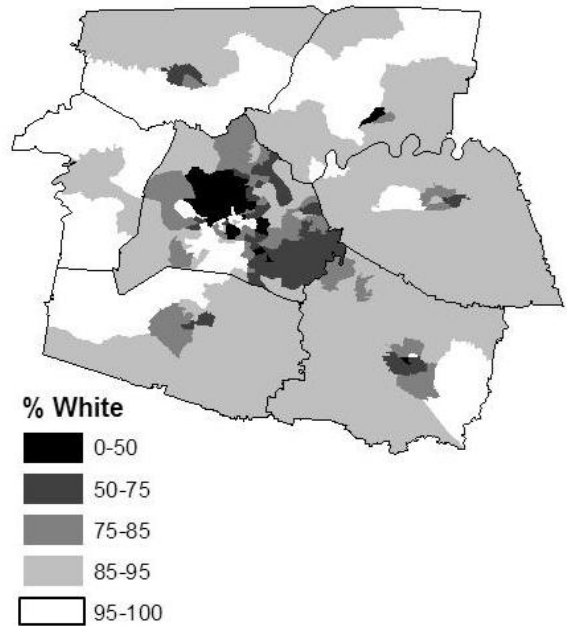
Map 4.2 Neighborhood Racial Composition
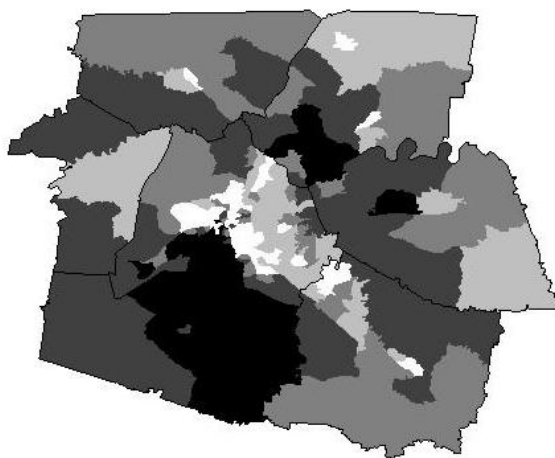
Map 4.3 Applicant Income (in Quintiles)
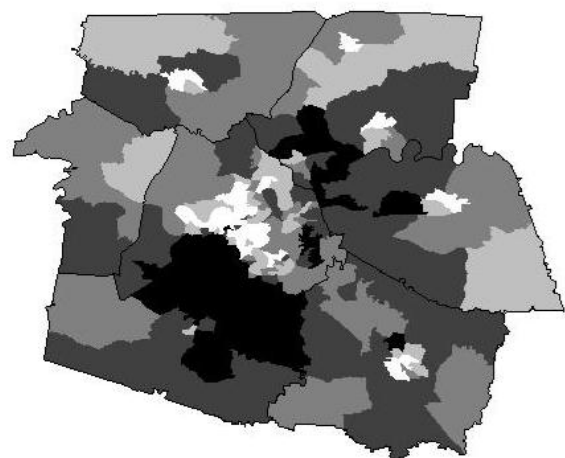
Map 4.4 Neighborhood SES (in Quintiles)

strongest tendency to favor minority applicants was in in the heart of North Nashville, the

historical center of Nashville's black community (see Map 4.2). Most of the rest of the

study area showed a slight preference for white applicants. Maps 4.3 and 4.4 show the

strong relationship between applicant and neighborhood socioeconomic characteristics.

Analysis

The analysis used is a spatial hierarchical linear model (Savitz & Raudenbush, 2009).

SHLM provides an elegant analytical tool for this research as it incorporates the

information provided by repeat observations within each tract as well as the spatial

relationships[15] between tracts. More detail on SHLM can be found in Chapter 3.

In the model building process, groups of variables will be added in the following

order:

1. Unconditional model (no variables)

2. Unconditional growth model (year variables)

3. Socioeconomic characteristics of applicants (income and loan amounts by race)

4. Neighborhood racial composition variables

5. Socioeconomic characteristics of neighborhoods

6. Changes in neighborhood racial composition

7. Changes in neighborhood socioeconomic status

---

[15] Specifically, the SHLM accounts for the correlation between error terms of adjacent tracts.

8. Discrete changes neighborhood racial composition

In its simplest form (without squared or interaction terms) the final model would look like this:

Level 1 equation:

$$\Delta\%W = \pi_{0n} + \pi_{1n}\, YEAR + \pi_{2n}\, WINC + \pi_{3n}\, MINC + \pi_{4n}\, WLOAN +$$
$$\pi_{5n}\, MLOAN + \pi_{6n} INCRAT + \pi_{7n}\, RATIO + \pi_{8n}\, ILW + \pi_{9n}\, ILM$$
$$+ \pi_{10n}\, ILR + \pi_{11n}\, CPW + \pi_{12n}\, CNSES + e_{jn}$$

Level 2 equations:

$$\pi_{0n} = \beta_{00} + \beta_{01}\, NW + \beta_{02}\, NS + r_0$$
$$\pi_{1n} = \beta_{10}$$
$$\pi_{2n} = \beta_{20}$$
$$\pi_{3n} = \beta_{30}$$
$$\pi_{4n} = \beta_{40}$$
$$\pi_{5n} = \beta_{50}$$
$$\pi_{6n} = \beta_{60}$$
$$\pi_{7n} = \beta_{70}$$
$$\pi_{8n} = \beta_{80}$$
$$\pi_{9n} = \beta_{90}$$
$$\pi_{10n} = \beta_{100}$$
$$\pi_{11n} = \beta_{110}$$
$$\pi_{12n} = \beta_{120}$$

Spatial dependence equation

$$r_0 = pW\, r_0 + b_0$$

The level-1 equation depicts how the dependent variable (differential success of white applicants versus minority applicants) is to be predicted by an intercept value (π_{0n}) and the year of application, the median white and minority incomes, the median white

and minority loan amounts, the white to minority income ratio, the white to minority loan

ratio, the loan to income (LTI) ratio of whites and minorities, and the ratio of these last

two LTIs, the change in neighborhood percent white since 2000, and the change in

neighborhood SES since 2000. The impact of each of these predictors is accounted for in

the parameter estimates of π_{1n} through π_{12n}.

As can be seen from the level-2 equations, the parameter estimates for π_{1n}

through π_{12n} will be the same for each tract. The intercept (π_{0n}), however, will be

estimated based on the neighborhood racial composition (NW) and neighborhood SES

(NS). This means that the predicted value for the intercept will potentially change for

each neighborhood in the sample. The spatial dependence equations shows how the level-

2 error on the intercept (r_0) will be decomposed into two components, neighborhood level

random error (b_0) and spatial dependent aspect of the data. In this equation, p is the

spatial correlation, W is the spatial weight matrix, and r_0 is a vector of random spatially

autoregressive effects.

<center>Results</center>

Model 1: Unconditional Model

This analysis aims to quantify the systematic impact of lending decisions on

neighborhood racial composition. Model 1 provides immediate clarification of the

connection between $\Delta\%W$ and neighborhood level factors. Since the intraclass

<center>75</center>

correlation (ICC) shows that most of the variation of Δ%W (91% in the non-spatial model, 95% in the spatial model) takes place at level-1, this means that less than 10% of the variation in the dependent variable is associated with neighborhood level of the analysis. There is, however, a strong spatial correlation in the outcome variable (0.86).

The results of the unconditional model also show that the intercept is 0.96, suggesting that, in the typical tract, the successful applicants are 1% more white than the overall percentage of white of applicants. Subsequent models attempt to account for this pattern through inclusion of time and financial variables at level-1 and to account for the smaller proportion of variance at level-2 by including neighborhood racial composition and neighborhood socioeconomic variables.

Model 2: Unconditional Growth Model

Time, but not time-squared, was associated with the outcome variable. The coefficient was negative, suggesting that while whites were favored across all years, the magnitude of this reduced across time. The model predicts that the Δ%W is a positive 1.14 in 2000 and this is reduced 0.79 in 2010. While time is a significant predictor, it only accounts for just over one-tenth of one percent of the total level-1 variance.

Table 4.2 Results of Models 1-3.

	Model 1		Model 2		Model 3	
	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
INTRCPT	0.960861	<0.001	0.966723	<0.001	0.951585	<0.001
YEAR			-0.03513	0.037	-0.02279	0.302
MEDWLOAN					-3.31847	0.095
MMLOAN					3.06749	0.118
RATIO					19.98707	0.035
ILW					14.28622	0.006
ILM					-13.4463	0.009
ILR					-17.7865	0.004
Level-1 Variance Explained	-		0.1%		1.1%	
Level-2 Variance Explained	-		1.5%		13.1%	
Spatial Correlation	0.86		0.86		0.85	

Model 3: Socioeconomic Characteristics of Applicants

The applicant economic characteristics were then added into the model. The income variables were not significantly associated with the dependent variable, but loan amounts and the loan-to-income ratios were. Perhaps surprisingly, this new model only accounted for marginally more of the level-1 variance: just over 1% of the total level 1 variance. Interestingly, it accounted for 13% of the level-2 variation (suggesting that there are consistent differences between tracts on these variables). However, that is only 13% of the 5% of the total variation associated with level-2 factors (i.e., 0.6% of total variance). The time variable also became insignificant in Model 3.

Model 4: Neighborhood racial composition

Both neighborhood racial composition variables (NW and NW2) were significant predictors of the dependent variable. The NW was positive and NW2 was negative, resulting in a pattern of predicting where banks favored white applicants in all

neighborhoods, but this effect peaked in 40% white neighborhoods and then declined as neighborhoods became more white (see Figure 4.4 below). As shown in Table 4.3, adding these variables accounts for approximately 40% of the total variation at level-2 and reduced the spatial correlation from 0.85 to 0.8.
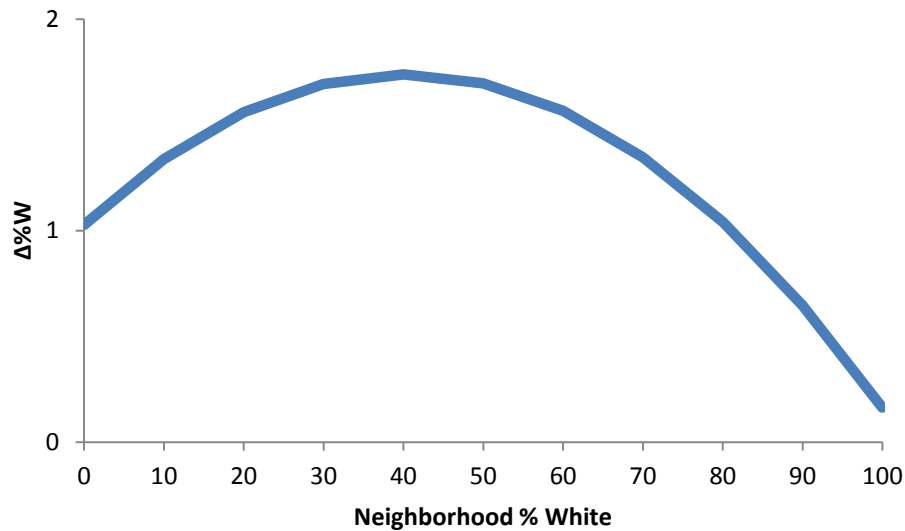
Figure 4.4 Relationship between neighborhood racial composition and Δ%W

Model 5: Neighborhood Socioeconomic Status

There were strong associations between the three neighborhood socioeconomic variables and the dependent variable, but NS2 and NS3 ceased to be significant when the neighborhood SES-Percent White (NWNS) interaction variable was included. The negative coefficient on NS suggests that as socioeconomic status increases, the differential in favor of whites decreases. On the other hand, the positive coefficient of the interaction term suggests that there is a stronger differential in favor of whites in

whiter-wealthier neighborhoods. This model explained 58% the level-2 variance and

reduced the spatial correlation coefficient to 0.77. Table 4.3 provides the results of Model

5.

Table 4.3 Results of Model 4, Model 5 and Final Model

		Model 4		Model 5		Final Model	
		Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
INTRCPT		1.028	0.014	-1.40082	0.038	-1.63273	0.016
	NW	0.035352	0.01	0.086351	<0.001	0.090388	<0.001
	NS			-1.23718	<0.001	-1.41788	<0.001
	NW2	-0.00044	<0.001	-0.00071	<0.001	-0.00072	<0.001
	NWNS			0.013056	<0.001	0.014655	<0.001
YEAR		-0.02909	0.183	-0.03268	0.134	-0.03501	0.108
MEDWLOAN		-3.55269	0.071	-3.1987	0.109	-3.05502	0.125
MMLOAN		3.805881	0.051	3.63181	0.062	3.627906	0.062
RATIO		23.1022	0.014	22.15246	0.019	21.99532	0.019
ILW		11.6417	0.026	12.352	0.018	12.43867	0.017
ILM		-11.8513	0.022	-12.531	0.015	-12.5976	0.014
ILR		-15.7259	0.011	-16.4802	0.007	-16.516	0.007
CNSES3						0.08071	0.016
Level-1 Variance Explained		0.8%		0.7%		0.9%	
Level-2 Variance Explained		40.4%		58.0%		59.5%	
Spatial Correlation		0.80		0.77		0.77	

Models 6, 7 and 8: Neighborhood Changes and Racial Composition Dummy Variables

Out of the nine variables added in final three models, only one was significantly associated with Δ%W (change in neighborhood SES-cubed [CNSES3]). None of the three change in neighborhood racial composition (CPW) variables predicted changes in the dependent variable. Changes in neighborhood SES at the more extreme ends seemed to be important since CNSES3 was significant, but the other two change in neighborhood SES variables (CNSES, CNSES2) were not. That is, large changes in the socioeconomic status of neighborhoods across the study period were positively associated with the differential favoring of white households. Similarly, the neighborhood dummy variables for minority, mixed, and traditional white neighborhoods were not significantly associated with the dependent variable. Adding the CNSES3 variable to the model did not have a substantial impact on the rest of the coefficients in the model and including this variable did not increase the amount of level-1 variation explained (see Table 4.3).

Figure 4.5 describes the predicted association between neighborhood racial composition and Δ%W based on the parameter estimates of the final model (see Table 4.3). The light grey line shows how Δ%W changes (the y-axis) when just looking at neighborhood racial composition (the x-axis) with all other predictors held at zero. It shows a tendency for lending decisions to make neighborhoods with very few white households slightly less white than would be the case if a racially proportionate amount of applicants were successful. For neighborhoods with populations of more than 20% white, the tendency is for these neighborhoods to be slightly whiter than would be predicted by the proportion of applicants who were white. This trend peaks at 65% white and then attenuates as it approaches 100% white neighborhoods. When the neighborhood socioeconomic status and its interaction with neighborhood racial composition is

included (the dark grey line) at their means for each neighborhood racial composition

grouping (0%-10% white, 10%-20% white, etc.), all neighborhoods are predicted to have

a higher proportion of successful applicants who are white than the proportion of total

applicants who are white. This effect is non-linear and declines in neighborhoods that are

greater than 80% white. The black line represents the predicted impact of lending

decisions when all variables are included at their means for each neighborhood racial

composition. Since the level-1 predictors were not highly associated with the dependent

variable we see very little difference between the black line and the dark grey line. On the

whole the final model accounted for just under 1% of level-1 variation, 59% of level-2

variation, and had a spatial correlation coefficient of 0.77.
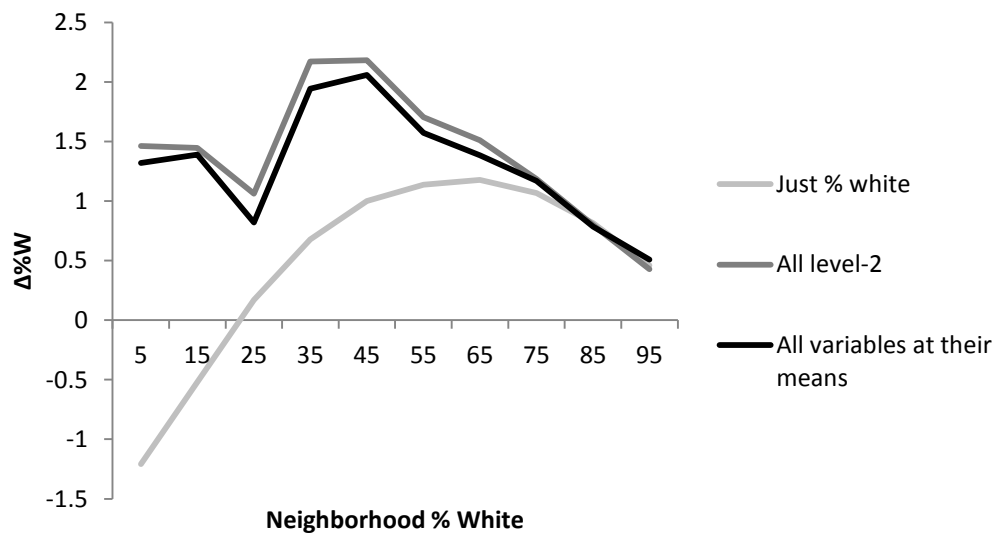
Figure 4.5 Relationship between neighborhood racial composition and Δ%W in the Final Model.

Table 4.4 shows the impact of lending decisions based on the average number of households applying for each tract based on its neighborhood racial composition. Assuming each variable at its average for each neighborhood composition (the same as the black line in Figure 4.5), the least white neighborhoods have an increase of white households that is equivalent to about one-half of a household. The maximum increase in white households is a bit over one household and occurs in the neighborhoods that are between 60 and 90 percent white. Since the average tract has around 2,000 households, this is equivalent to a relatively minor increase of (1.25/2000 = 0.06%) in neighborhood percent white. If this pattern was maintained for 10 years, that would translate to lending decisions making a neighborhood just over half of one percent (0.63%) whiter than it would have been if applicants were approved in a racially proportionate way. As will be noted below, further research is needed to tease out this complex set of findings.

Table 4.4 Impact of lending decisions on in-movers by neighborhood racial composition

Tract % White	Avg # Apps	Avg # White Apps	Avg # Approved Apps	% of Apps who were White	Predicted % of Successful Apps who were White	Increase in # of White Hhlds in Tract
0-10	45	11	30	24%	26%	0.4
10-20	50	22	38	44%	46%	0.5
20-30	83	37	66	45%	46%	0.5
30-40	63	43	50	69%	70%	1.0
40-50	67	47	52	70%	72%	1.1
50-60	85	54	69	63%	65%	1.1
60-70	117	80	97	69%	70%	1.3
70-80	119	96	101	80%	82%	1.2
80-90	183	157	160	86%	86%	1.3
90-100	169	158	149	93%	94%	0.8

Discussion

This study examined whether lenders systematically altered the racial composition of neighborhoods by the aggregate impact of their lending decisions. To do this it looked at the difference between the racial composition of successful loan applicants and the racial composition of the total pool of applicants for each tract in each year. Early analyses showed that on average, lending decisions made the successful applicants about one percent whiter than the original pool of applicants for a particular tract. The unconditional model also showed that most of the variation in the change in the racial composition of applicants was not associated with neighborhood level factors. That is, knowing which neighborhood a household was applying in only helped to predict a small amount of the variation in the dependent variable.

Nevertheless, there were statistically significant relationships between neighborhood level characteristics and the dependent variable. As was shown in Table 4.4, 60% to 90% white neighborhoods had the most substantial increase in the number of white households who would be moving into a tract (approximately 1.25 households when all other variables were included at their means for that neighborhood racial composition).

It is clear that lenders are not solely making their lending decisions based on neighborhood racial composition (in fact, well over 90% of their decision making is based on factors not related to neighborhood-level factors), it is also clear that there is a relationship between the neighborhood and spatial factors and the likelihood that white

applicants will be approved at a higher rate than non-white applicants. When the results are examined in terms of the differential approval of white households, 30% to 50% white neighborhoods are where there are the highest differentials in approval rates. However, as noted above, when data is examined in terms of increase in actual number of white households moving into a neighborhood, majority white neighborhoods (60%-90% white) are where we see the greatest impact of lending decisions (this is due to the larger number of applications in these neighborhoods). This impact does not seem to be mediated through the socioeconomic status of the neighborhood, and the relationship between neighborhood racial composition and differential success of white applicants occurs after many economic controls are included about the composition of applicants. This pattern of results leaves open the question of whether lenders are participating in geographic contingent lending – favoring white applicants in whiter neighborhoods – a topic to be explored in Chapter 5.

The pattern of results also shows that a closer analysis of which neighborhoods are seeing the strongest impact on Δ%W may be highly productive. The combination of statistically significant association of neighborhood percentage white, high spatial dependence, and a low proportion of variance associated with level-2, suggests that perhaps some isolated neighborhoods in certain parts of the metropolitan area are seeing a distinct pattern of lending decisions, while other neighborhoods do not have a systematic pattern in Δ%W. A highly specific pattern will not necessarily be revealed by a multiyear study of multiple neighborhoods such as this one, and a mixed method exploration of both statistical patterns and institutional approaches to certain parts of the

metropolitan area would be better suited to enhancing understanding of these potentially

highly localized patterns.

CHAPTER V

DO LENDING INSTITUTIONS DISCRIMINATE ON THE BASIS OF APPLICANT

RACE, NEIGHBORHOOD RACIAL COMPOSITION, OR, THE INTERACTION

BETWEEN THEM?

While analysis 2 simply investigated whether there may be factors other than
preferences (i.e,, lending decisions) playing a role in segregation, question 3 asks whether
there is evidence of discrimination in lending decisions and whether this, specifically, is
contributing to racial residential segregation. To this end, this question explores three
types of discrimination: discrimination based on race of applicant, discrimination based
on racial composition of neighborhood (redlining), and discrimination based on the
interaction of applicant and neighborhood racial characteristics. The third form of
discrimination, defined by Holloway (1998) as "geographically contingent lending",
could provide a direct pathway from discriminatory lending decisions to segregation. In
this chapter, I expand on this idea and look at the findings from the most important
studies in the area, present the hypotheses, and then introduce the models that will be
used to test the proposed hypotheses.

Research on Racial Discrimination in Lending

Historically, mortgage discrimination researchers focused on either discrimination against minority neighborhoods (redlining) or discrimination against minority individuals (Holloway, 1998). In many ways, the difference between the two types of discrimination was minor – in America's strongly segregated cities, minority applicants generally applied for loans in minority neighborhoods. Increasingly, however, minority households are searching for housing outside of traditional minority neighborhoods.

The most comprehensive and controversial lending discrimination study is known as the Boston Fed Study (Munnell, Geoffrey, Lynn, & James, 1996). The study focused on discrimination based on the race of applicant. The original findings, using data from 1990, found large differences in the denial rates between whites and blacks, even after controlling for nearly every important criterion used in loan under-writing. The study found that minority applicants were eight percent more likely to be denied a loan in comparison to similarly qualified whites (Munnell, et al., 1996). Nevertheless, the study came under strong criticism, both methodologically and conceptually. A decade of reanalysis of the same dataset, however, has come to the conclusion that there is strong evidence that lenders in Boston engaged in racial discrimination in lending (Ross & Yinger, 2002). The Boston Fed Study has not been replicated because, to date, no lenders have provided the refined level of data that was utilized in that study. In a study also using the Boston Fed data, Tootell (1996) looked into redlining in the Boston area and found no relationship between neighborhood racial composition and denial probability.

Two studies by Steven Holloway have looked more closely at the interaction between race of applicants and the racial composition of the neighborhood they were applying to live in. In his study of Columbus, Ohio – based on 1992 HMDA data – Holloway (1998) uses multilevel modeling to analyze mortgage lending decisions by census tract. After controlling for socioeconomic factors, he tests whether there is evidence of discrimination against minority applicants. He finds that, on average, there is no evidence of discrimination. However, the variance of the census tract-level error term on race is significantly greater than zero, suggesting that the importance of race changes across census tracts. Holloway then tests for the importance of redlining by examining whether tract racial composition is predictive of loan denial. His results, like those of others (Abariotes et al., 1993; Tootell, 1996), suggest that redlining was not an issue. However, when he tests for an interaction between neighborhood racial composition and race of applicant he finds that there is a strong effect, and that minority applicants moving into white neighborhoods, particularly those with large loans, are denied at rates significantly greater than chance. Similarly, white applicants in minority neighborhoods, particularly those applying for smaller loans, were more likely to be denied than minority applicants.

Holloway's Columbus study was replicated by Holloway and Wyly (2001) using 1996 HMDA data from Atlanta. In this study, unlike Holloway (1998), the authors found that black applicants were more likely to be denied in all tracts. However, like the earlier study, they found no evidence of redlining and strong evidence of geographically contingent lending – that is, lending decisions were not affected by the racial composition of the neighborhood, but there was a significant interaction between the race of applicant

and neighborhood racial composition. Specifically, they found that in Atlanta black households applying for loans in white affluent neighborhoods were much more likely to be denied. These results echo those found in Chapter 4: lending decisions seem to favor whites most in white neighborhoods.

Analysis 3 builds on the lending discrimination research of Holloway (1998) and Holloway and Wyly (2001). It tests three hypotheses:

H1 - Individual Discrimination: Lenders are more likely to approve the loans for white applicants than non-white applicants, even after controlling for key socioeconomic factors.

H2 - Redlining: Lenders are more likely to deny loans that are for houses in minority neighborhoods, even after controlling for socioeconomic factors.

H3 - Geographically Contingent Lending: Lenders are more likely to deny loans to minority applicants attempting to move to white neighborhoods, even after controlling for socioeconomic factors.

I will use a spatial 2-level logistic model to predict likelihood of denial based on individual characteristics (e.g., race, gender, income, probability of credit score denial), loan and lender characteristics (loan amount, loan to income ratio, type of lending institution), and neighborhood characteristics (socioeconomic status [income, education, housing value], vacancy rate, population change, neighborhood racial composition). This will allow for the testing of the three hypothetical forms of discrimination: traditional (minority applicants are denied loans), redlining (minority neighborhoods are denied

loans), and geographically contingent (denial probability is dependent on an interaction between race of applicant and neighborhood racial composition). To test for geographically contingent lending denials, a cross level interaction where the influence of applicant race is predicted by neighborhood level variables will be modeled (more detail on these equations is provided below).

In combination, this third set of analyses will help answer the question about whether mortgage lenders appear to be discriminating on the basis of applicant race, neighborhood composition, and/or the interaction between the two.

Data

From the Nashville MSA HMDA Loan Application Register I took the individual loan records for valid census tracts in the seven county study area where the following were true [cases remaining]:

1. It was a home purchase loan [593,387]

1. The borrower was going to reside in the home [539,884]

2. Loan was either approved or denied [367,107]

2. Applicant race was reported [320,718]

3. Case was not flagged as edited (indicative of an incomplete record) [268,053]

4. Inflation adjusted income and loan amounts were between $20,000 and $500,000 [255,184]

Using this data set, I created the following variables:

- **APP:** Dummy variable indicating whether the loan was approved or denied (the dependent variable)

- Agency dummies representing each different type of lending institution (i.e., those regulated by the Office of the Comptroller of the Currency **(OCC)**, Federal Reserve System **(FRS)**, Federal Deposit Insurance Corporation **(FDIC)**, Office of Thrift Supervision **(OTS)**, National Credit Union Administration **(NCUA)**, or other)

- **LCON:** Conventional loan dummy (as opposed to an Federal Housing Agency [FHA], Veterans Administreation [VA], or Rural Development [RD] backed loan)

- **LINC & LLOAN:** Natural log of inflation adjusted income and loan amounts

- **LTI:** Loan to income ratio variable

- **PNHW & CNHW:** Non-Hispanic white applicant and co-applicant dummies

- **PF:** Female primary applicant, no co-applicant dummy

- Denied for credit history dummy (if any of the three denial reasons involved credit history), to help create the instrumental variable quantifying the probability that an applicant would be denied due to their credit history (**FV**).

Most of the census tract variables used in Holloway and Wyly[16] were strongly correlated with the NSES variable I had created for the earlier analyses. The two exceptions were population change and vacancy rate. In order to simplify comparisons

---

[16] Percent white, Percent with bachelor degrees, Median income, Median value, Change in Population, Rent to value, Vacancy rate, Percent of Population that did not move in last 5 years.

across models, I opted to use the NSES index variable in conjunction with the population change and vacancy rate variables in my models.

After univariate analysis of each variable was carried out, I dropped observations that had outlying cases. This included less than 5% of the total cases in the 2000-2004 data (5,903 cases) and less than 2.5% of the 2005-2009 cases (2624 cases removed)[17].

Analyses

Due to the large sample size and complexity of the analysis, I split the sample into three loan cohorts. The first group included loan applications between 2000 and 2002, when all applications were included located within 1990 census tracts. The second group included applications between 2003 and 2004, when there was inconsistent usage of 1990 and 2000 census geographies by HMDA reporting institutions. The third cohort included applications between 2005 and 2009 (all using Census 2000 tracts). Neighborhood-level predictors were based on Census 2000 data for the first two cohorts, while 2005-2009 ACS data was used as the neighborhood predictors for the third data set (the variables were the same, however).

For each of the three groups, I followed Holloway and Wyly's (2001) method of creating an instrumental variable predicting the likelihood of an applicant having bad credit. This variable attempts to provide a control variable for credit history, since this is

---

[17] The larger proportion of the earlier cohorts that were outliers was probably due to the process of merging 2000 census characteristics into 1990 tracts that was described in Chapter 3 (this was unnecessary in the later sample as the 2005-2009 cases were all listed in 2000 census tracts).

not included in HMDA data (however, denial reasons are an optional field in HMDA reports). The variable was created by running a logistic regression predicting the likelihood of denial due to poor credit history on half the applications of each loan cohort and then using the coefficients on each variable to predict the probability of poor credit history on the second half of the group (the data that was used in the subsequent models). Correlations between the fitted value and outcome were consistent between each half of each loan cohort, suggesting that the model was effective at identifying credit history denials in the second data sets. Histograms and descriptive statistics were also compared for each group and found to be consistent across cohorts and across each half of the data.

The descriptive statistics for each variable used in the spatially dependent logistic multilevel models can be found below in Table 5.1.

Table 5.1

	2000-2002					2003-2004					2005-2009				
	N	MEAN	SD	MIN	MAX	N	MEAN	SD	MIN	MAX	N	MEAN	SD	MIN	MAX
Dependent Variable (Loan Approved dummy variable)															
APP	34371	0.85	0.36	0	1	25587	0.87	0.34	0	1	57629	0.88	0.32	0	1
Type of Lending Institution (Regulating Agency dummies)															
OCC	34371	0.27	0.44	0	1	25587	0.27	0.44	0	1	57629	0.2	0.4	0	1
FRS	34371	0.17	0.38	0	1	25587	0.27	0.44	0	1	57629	0.26	0.44	0	1
FDIC	34371	0.05	0.21	0	1	25587	0.05	0.22	0	1	57629	0.12	0.32	0	1
OTS	34371	0.07	0.25	0	1	25587	0.03	0.16	0	1	57629	0.03	0.18	0	1
NCUA	34371	0	0.07	0	1	25587	0.01	0.07	0	1	57629	0.01	0.1	0	1
Loan Type (Conventional Loan Dummy)															
LCON	34371	0.67	0.47	0	1	25587	0.74	0.44	0	1	57629	0.77	0.42	0	1
Income and Loan Amounts (Natural Log of Income and Loan Amounts, and Loan to Income Ratio)															
LINC	34371	4.14	0.53	3.01	6.21	25587	4.12	0.54	3.02	6.21	57629	4.15	0.55	3	6.21
LLOAN	34371	4.79	0.6	3.01	6.21	25587	4.86	0.6	3.02	6.21	57629	4.86	0.66	3	6.21
LTI	34371	1.17	0.15	0.55	1.89	25587	1.19	0.15	0.5	1.75	57629	1.18	0.17	0.53	1.85
Applicant Race and Gender (Applicant White, Co-applicant White, Primary Applicant Female with no Co-applicant)															
PNHW	34371	0.84	0.36	0	1	25587	0.82	0.38	0	1	57629	0.81	0.39	0	1
CNHW	34371	0.43	0.5	0	1	25587	0.38	0.49	0	1	57629	0.35	0.48	0	1
PF	34371	0.22	0.42	0	1	25587	0.25	0.43	0	1	57629	0.26	0.44	0	1
Year Variables[18]															
Y1	34371	0.32	0.47	0	1	25587	0.46	0.5	0	1	57629	-0.43	1.3	-2	2
Y2	34371	0.32	0.46	0	1						57629	1.87	1.66	0	4
Y3											57629	-1.2	4.73	-8	8
Credit History Instrument (Predicted Likelihood of Denial Due to Poor Credit History)															
FV	34371	0.05	0.05	0	0.3	25587	0.04	0.03	0	0.3	57629	0.03	0.02	0	0.3

---

[18] These are dummy variables in the first two loan cohorts (2000-2002,2003-2004), and linear, squared and cubed terms in 2005-2009 data.

| | 2000-2002 | | | | | 2003-2004 | | | | | 2005-2009 | | | |
	N	MEAN	SD	MIN	MAX	N	MEAN	SD	MIN	MAX	N	MEAN	SD	MIN	MAX
Level 2 Variables (Tract % White, Tract SES, (Tract % White)^2, (Tract SES)^2, Vacancy Rate, Population Change, Race x SES Interaction)															
NW	166	76.04	24.94	0.71	97.73	227	75.92	24.41	0.71	98.49	215	70.64	25.25	0	98.78
NS	166	-0.1	1.43	-2.76	4.95	227	0.1	1.61	-2.76	5.37	215	-0.02	1.6	-2.39	5.59
NW2	166	6400.92	2835.04	0.51	9550.19	227	6357.13	2827	0.51	9701.11	215	5624.54	2930.66	0	9758.2
NS2	166	0.31	4.18	-7.6	24.54	227	1.12	5.44	-7.6	28.82	215	2.54	4.67	0	31.25
VAC	166	5.9	2.37	1.94	14.83	227	5.63	2.33	1.09	14.83	215	0.08	0.05	0	0.28
POPC	166	1.14	0.21	0.63	2.11	227	1.15	0.21	0.63	2.35	215	1.15	0.22	0.63	2.35
NWNS	166	13.01	112.87	-167.68	457.85	227	29.13	133.29	-167.68	520.36	215	22.35	125.41	-175.29	539.56

Maps 5.1 and 5.2 show the spatial distribution of each of the variables used in the

analyses. The maps show each variable broken out into quintiles (5 equally sized

groups), with tracts scoring in the lowest quintile of each variable colored white and the

shading of tracts getting progressively darker as the values increase (the green census

tracts demark areas of no data). The first map shows the average approval rates for each

census tract, with the highest approval rates found in the south-western edge of Davidson

County and the neighboring portion of Williamson County. The second map shows the

proportion of applicants who are white in each census tract. Not surprisingly, the ring

counties have the highest proportion of white applicants. Of the lending institution maps,

we see that Federal Reserve (FRS) and Office of the Comptroller of the Currency (OCC)

regulated lenders are highly active in the wealthier areas of the metropolitan area (e.g.,

Williamson County). Similarly, these wealthier areas are more likely to utilize

conventional loan products. Lenders who were not overseen by a major regulator were

most active in the north Nashville and Antioch areas – two areas that have much higher

concentrations of minorities. Davidson County census tracts have much higher numbers

of primary applicants who are female with no coapplicants. The northwestern quadrant of

the metropolitan area has higher rates of denial due to credit history (including North

Nashville, but also incorporating Sumner County).

Overall, the maps show clear neighborhood, county and regional dynamics.

Southern Davidson County and Williamson County often have similar values on the

different variables mapped. Similarly, central and southeastern Davidson County tend to

have their own distinct patterns as does the northern horseshoe of counties: Cheatham,

Robertson, Sumner, and Wilson counties. Rutherford County, to the southeast of the

metropolitan area, tends to have a set of patterns that fluctuate between those of central

and southeastern Davidson County, and the northern horseshoe of counties.
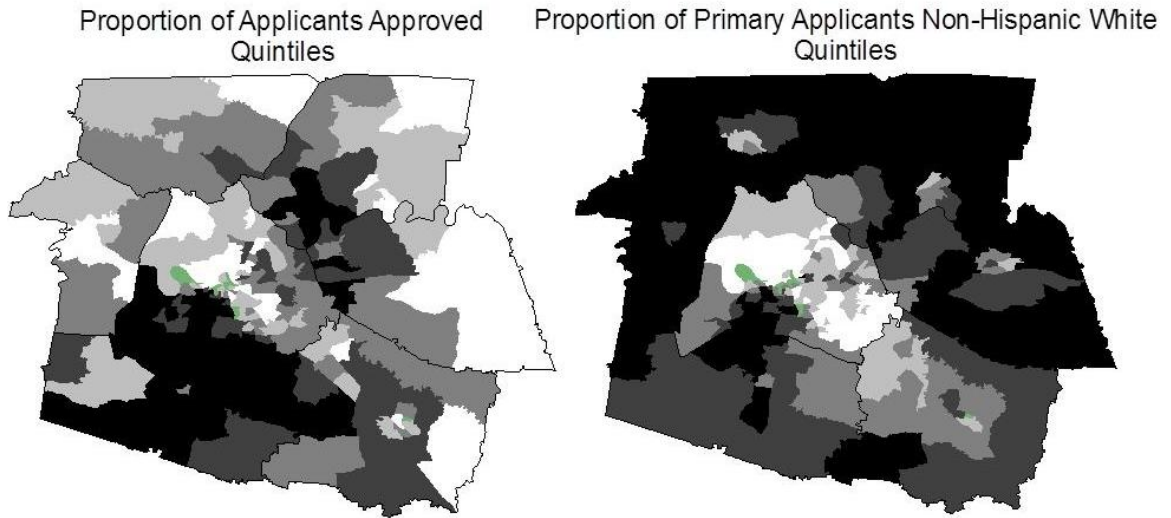
Maps 5.1 Level-1 Variables

Proportion of Applicants using Conventional Loans
Quintiles
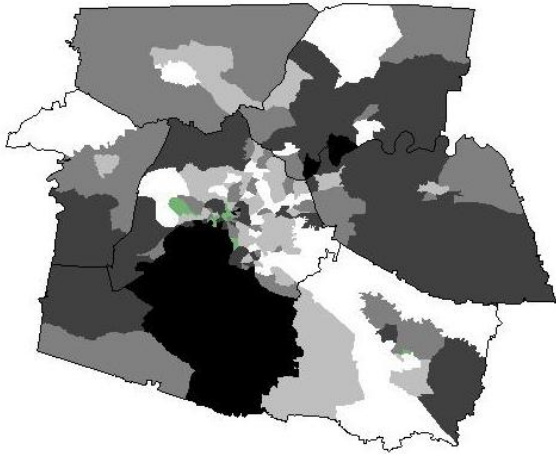
Proportion of Applicants using OCC Lenders
Quintiles
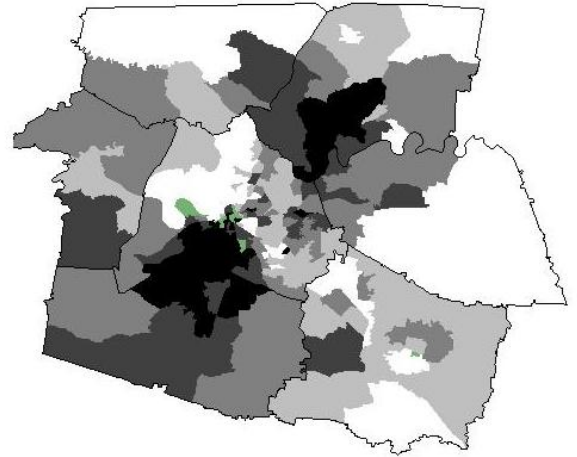
Proportion of Applicants using FDIC Lenders
Quintiles
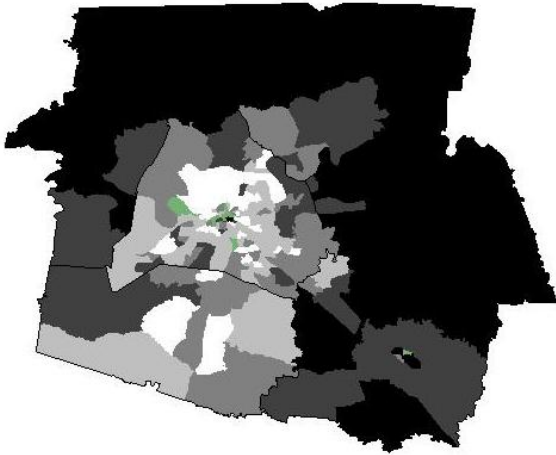
Proportion of Applicants using FRS Lenders
Quintiles
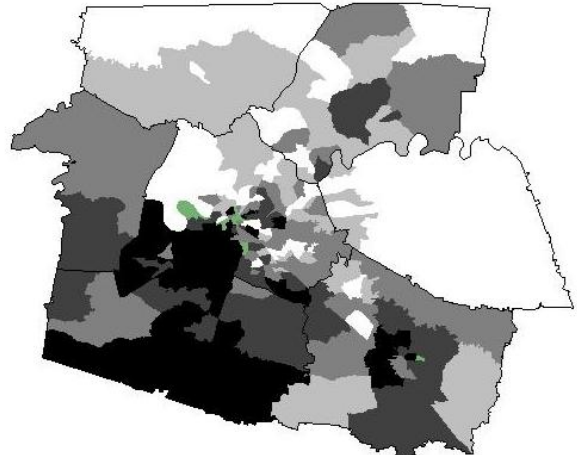
Proportion of Applicants using OTS Lenders
Quintiles
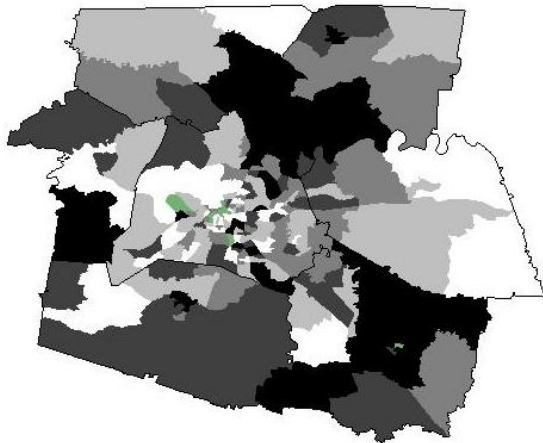
Proportion of Applicants using Other Lenders
Quintiles
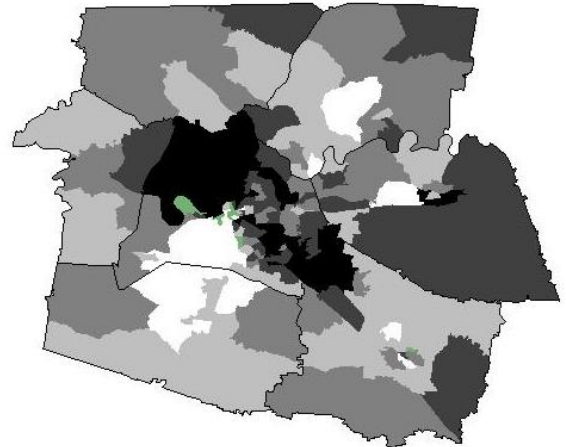
Maps 5.1 Level-1 Variables (continued)

Proportion of Primary Applicants Female, No Co-Applicant
Quintiles
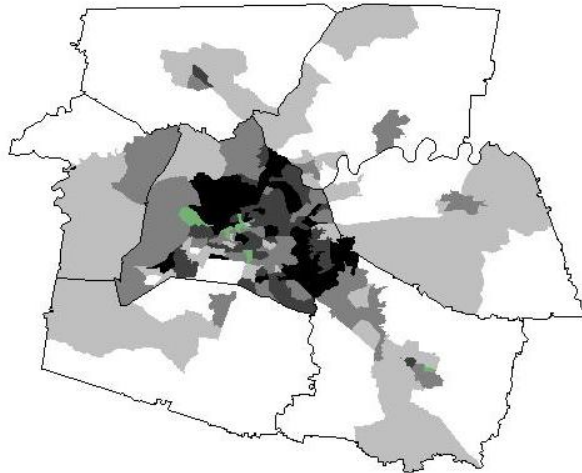
Average Loan to Income Ratio
Quintiles
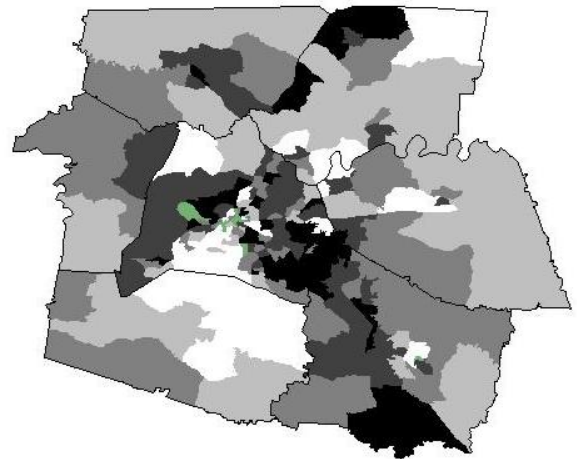
Proportion of Co-Applicants Non-Hispanic White
Quintiles
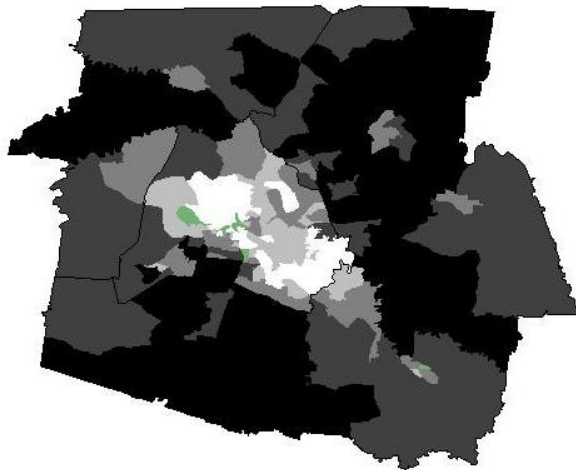
Average Loan Amount (Natural Log)
Quintiles
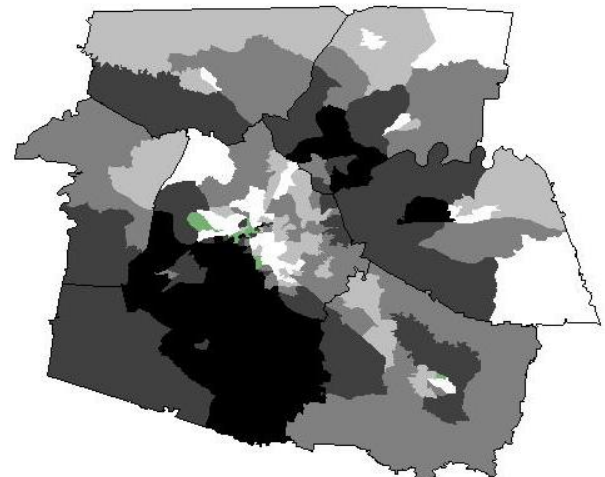
Average Income (Natural Log)
Quintiles
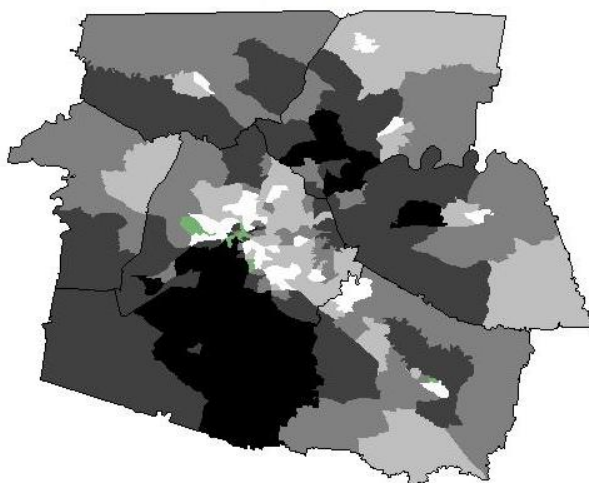
Proportion of Applicants Denied Due to Credit History
Quintiles

Maps 5.1 Level-1 Variables (continued)

99
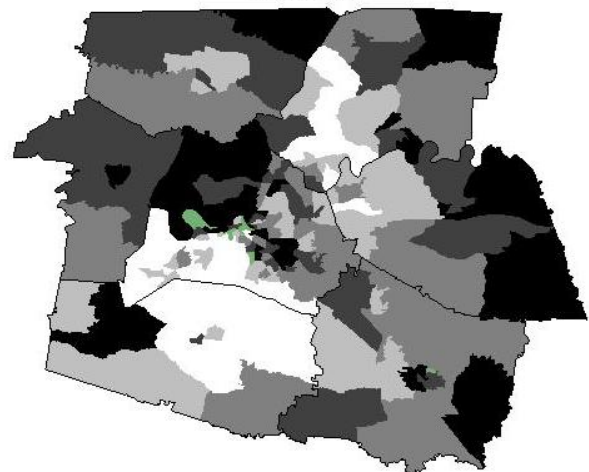
Neighborhood Percent White x Neighborhood SES Interaction
Quintiles
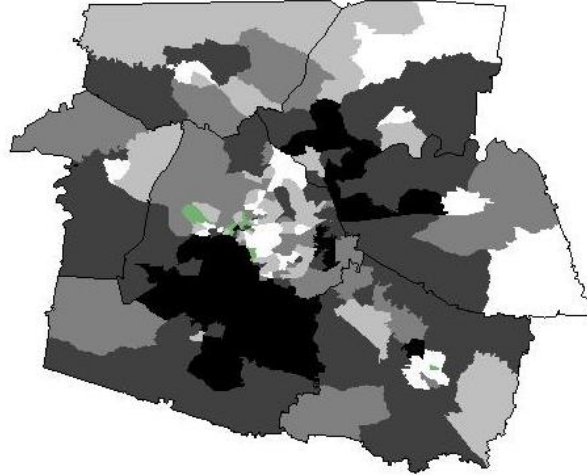
Neighborhood Percent White
Quintiles
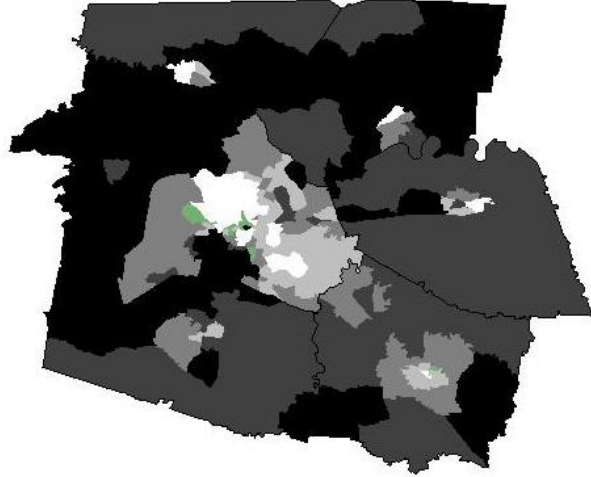
Neighborhood SES
Quintiles
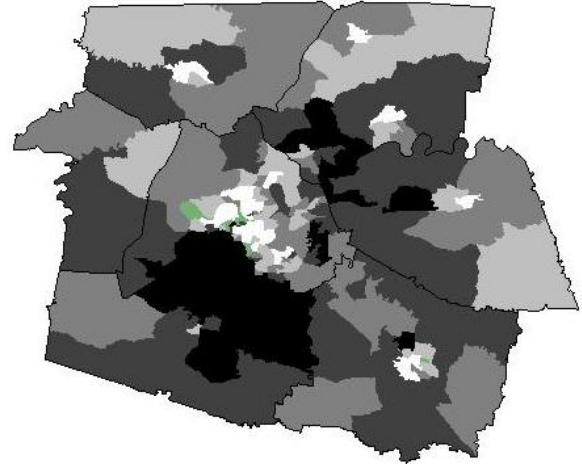
Population Change
Quintiles
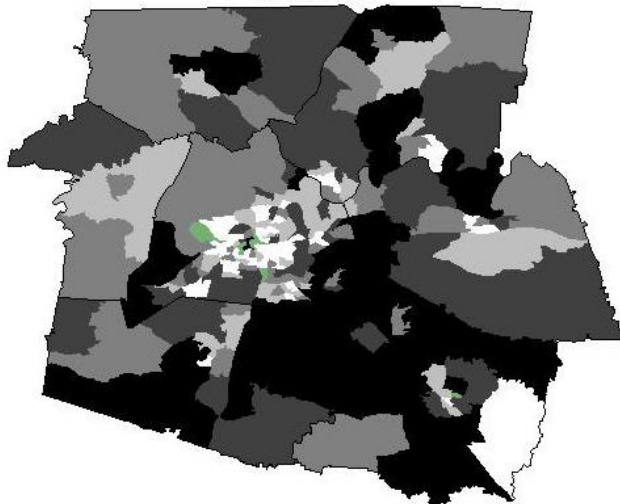
Vacancy Rate
Quintiles
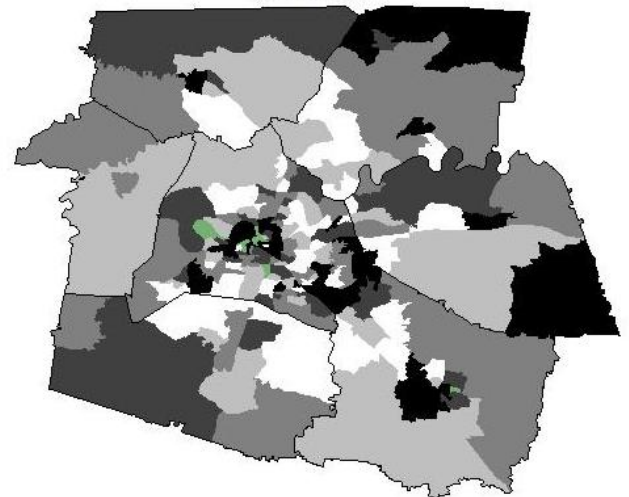
Maps 5.2 Level-2 Variables

In order to ensure comparability between the three loan cohorts, I began with the same, relatively comprehensive, predictive model testing H1 and H2 (individual discrimination and redlining, respectively). I refer to this model as the comparison model (or M1). This model is described in the equations below. After this model was estimated for each loan cohort, I then explored whether a better model could be estimated by including an interaction term between neighborhood racial composition and neighborhood SES, neighborhood vacancy rates, and neighborhood population change. I refer to this second model as the expanded model (or M2). Following that, I examined whether there was evidence of geographic contingent lending (H3) for each subsample. This third model is referred to as the geographic contingent model (or M3).

Level-1 Equation for the Comparison Model:

Log (Probability [APP=1] / (1- Probability [APP=1]))

$$= \pi_{0n} + \pi_{1n} \text{LCON} + \pi_{2n} \text{LINC} + \pi_{3n} \text{LLOAN} + \pi_{4n} \text{PNHW} + \pi_{5n} \text{CNHW} +$$

$$\pi_{6n} \text{LTI} + \pi_{7n} \text{FV} + \Sigma\Upsilon_{1i}(\text{INSTITUTIONTYPE}_i) + \Sigma\Upsilon_{2i}(\text{YEAR}_i) + e_{pn}$$

π_{0n}	$=$	$\beta_{00} + \beta_{01}\text{NW} + \beta_{02}\text{NW2} + \beta_{03}\text{NS} + \beta_{04}\text{NS2} + r_0$
π_{1n}	$=$	β_{10}
π_{2n}	$=$	β_{20}
π_{3n}	$=$	β_{30}
π_{4n}	$=$	β_{40}
π_{5n}	$=$	β_{50}
π_{6n}	$=$	β_{60}
π_{7n}	$=$	β_{70}
π_{8n}	$=$	β_{80}
π_{9n}	$=$	β_{90}
π_{10n}	$=$	β_{100}
π_{11n}	$=$	β_{110}
π_{12n}	$=$	β_{120}
π_{13n}	$=$	β_{130}
π_{14n}	$=$	β_{140}
π_{15n}	$=$	β_{150}

Spatial dependence equation

$$r_0 = pW\, r_0 + b_0$$

The level-1 equation shows how the logit of approval is predicted by the intercept (π_{0n}) and the loan type, the income and loan amounts, the race of applicant and coapplicant, the loan to income ratio, the likelihood of denial due to credit history, the type of lending institution and the year of application. As with the other independent variables in earlier models, the parameters are estimated to have a consistent impact in all

tracts. The value of the intercept, however, is dependent on the neighborhood racial composition and neighborhood SES (as well as their squared terms).

Results

Comparison Model (M1)

The results of this model are displayed in Table 5.2 below. The results for each analyses showed a strong positive impact associated with non-Hispanic white primary applicants (PNHW). This demonstrates that, controlling for the neighborhood, loan and other applicant characteristics in the model, white applicants were favored over non-white applicants (H1 is strongly supported). Figure 5.1 shows these relationships, with the dotted lines (white applicants) always being higher than the solid lines (minority applicants). The results also showed a non-linear relationship between neighborhood percentage white and applicant approval, although the 2000-2002 data coefficient is slightly below the significance threshold on the non-squared term. As can be seen in Figure 5.1, the impact of neighborhood percent white peaks around 50% white. Thus, there is inconsistent evidence for redlining (H2) when looking at the results of all three models. On the whole, these results show that lending institutions are more likely to approve white applicants and are more likely to approve applicants who want to live in racially mixed neighborhoods, even after controlling for neighborhood SES.

Table 5.2 Results of Comparison Model (M1) for each Loan Cohort

Fixed Effect	2000-2002		2003-2004		2005-2009	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
INTRCPT2	1.447532	<0.001	1.413116	<0.001	1.216718	<0.001
PW	0.011219	0.099	0.012577	0.037	0.020898	<0.001
NS	0.366743	<0.001	0.210544	<0.001	0.201083	<0.001
PW2	-0.00013	0.022	-0.00013	0.01	-0.00019	<0.001
NS2	-0.0602	<0.001	-0.03192	0.025	-0.01767	0.036
OCC	0.715973	<0.001	0.391326	<0.001	0.156496	<0.001
FRS	0.743489	<0.001	0.770828	<0.001	0.348338	<0.001
FDIC	0.90124	<0.001	0.945146	<0.001	0.937626	<0.001
OTS	0.922876	<0.001	0.205978	0.09	0.200797	0.009
NCUA	0.553337	0.018	0.543571	0.046	0.112897	0.468
LCON	-0.56269	<0.001	-0.31989	<0.001	-0.29271	<0.001
LINC	-0.31064	0.244	-0.56627	0.032	-1.05385	<0.001
LLOAN	1.014926	<0.001	1.028786	<0.001	1.268269	<0.001
PNHW	0.40339	<0.001	0.354351	<0.001	0.472422	<0.001
CNHW	0.063009	0.127	0.078171	0.121	0.231822	<0.001
PF	0.164668	<0.001	0.029682	0.546	0.082304	0.013
LTI	-3.16075	<0.001	-3.45914	<0.001	-4.95255	<0.001
Y	0.18014	<0.001	0.080051	0.055	-0.11611	<0.001
Y2	0.082271	0.04	-	-	0.007903	0.394
Y3	-	-	-	-	0.032321	<0.001
FV	-4.29403	<0.001	-5.8573	<0.001	-3.11591	0.007

Table 5.3 shows the estimated intraclass correlation (ICC) assuming that level-1 variance is $\pi^2/3$ (Raudenbush & Bryk, 2004). It also shows the percentage of the level-2 variance explained by the models, the spatial correlation in the error terms and the change in spatial correlation from the unconditional model. The estimated ICC shows, for all three cohorts, the majority of variation in loan approval decisions is due to applicant level factors rather than neighborhood level factors. This should be expected, as lenders are likely to be more concerned with the credit worthiness of applicants than with the neighborhood that they plan on moving in to. In the 2000-2004 loan cohorts, the models helped explain two-thirds to three quarters of the level-2 variation, and also reduced the

spatial dependence of the error terms. A smaller amount of level-2 variation was

explained in 2005-2009 model, but the spatial dependence in the model was reduced from

correlation coefficient of almost 0.9 to 0.32. This reduction in spatial correlation suggests

that what the unconditional model was calculating as spatial dependence was in fact

accounted for by spatially correlated values in tract percentage white and SES.

Table 5.3 Model Fit for Comparison Models (M1)

	2000-2002	2003-2004	2005-2009
Estimated Intraclass Correlation (% of Variance associated with Level-2)	6.2%	3.5%	2.9%
Percentage of Level 2 Variance Explained	75.7%	67.5%	42.4%
Spatial Correlation in Dependent Variable	0.679	0.524	0.323
Change in Spatial Correlation from Unconditional Model	-0.134	-0.357	-0.549
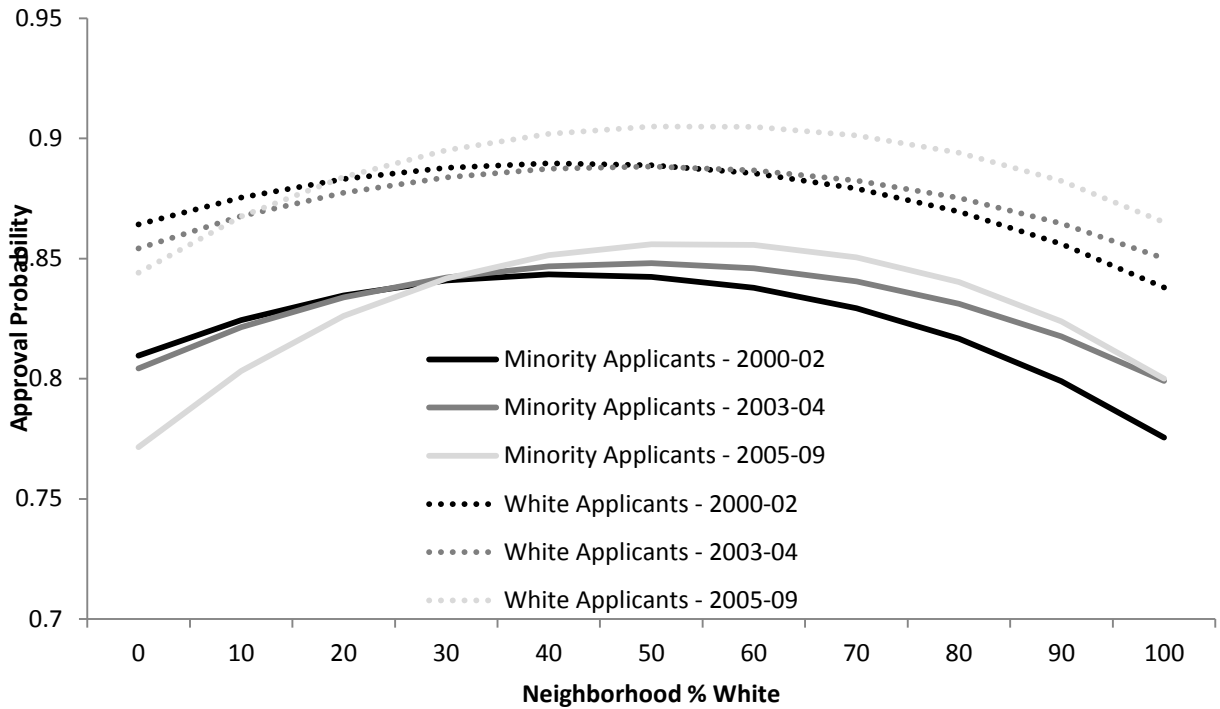
Figure 5.1 Relationship between neighborhood racial composition and probability of approval for white and minority applicants for each loan cohort (M1)

2000-2002 Loan Cohort – M2 – Expanded Model

Model 1 did not include three variables previous research suggested potentially impact the likelihood of denial: vacancy rates, population change and an interaction term between neighborhood racial composition and neighborhood SES. When these neighborhood level variables were added to the comparison model (M1), it was found that vacancy rates and population change variables were not significant predictors of denial rates, but the interaction between neighborhood racial composition and socioeconomic status was positively associated with denial rates. The positive coefficient on the interaction term suggests that there was a compounding effect of higher

socioeconomic status and neighborhood percent white: applications were increasingly likely to be approved in whiter-wealthier neighborhoods than would be predicted by just examining the coefficients on each of these variables independently. The full results of this model can be seen in Table 5.4.

The comparison of the M1 and M2 curves in Figure 5.2 depicts how applicants in high minority neighborhoods were disadvantaged compared to all other applicants. In fact, a minority applicant in zero percent white neighborhood has 0.08 lower probability of becoming approved compared to a minority applicant in a 100% white neighborhood (40%-80% white neighborhoods were the most favored).

2000-2002 Loan Cohort – M3 – Geographic Contingent Lending

M3 built on the results of M2 to assess whether there was evidence of neighborhood contingent lending in the 2000-2002 data (H3). The neighborhood contingent lending models have a similar specification to the ones above, but the applicant race parameter (π_{4n}) is predicted by neighborhood level variables. For example,

$$\pi_{4n} \quad = \quad \beta_{40} + \beta_{41}NW \ + \beta_{42}NS$$

That is, the parameter estimate for applicant race is predicted by a level-2 intercept (β_{40}; the average impact of being white after neighborhood racial composition and SES has been taken into account), neighborhood racial composition and neighborhood SES. If β_{41} and β_{42} were non-significant, it would show that there was no evidence of geographic contingent lending.

The 2000-2002 loan cohort had strong evidence of neighborhood-level variables interacting with the race of applicant (see Table 5.4). As the light gray lines in Figure 5.2 show, white applicants were less likely to be approved in high minority neighborhoods (10% white or less), but more likely to be approved in all others. As can be seen by the shape and position of the two grey dotted lines, M2 and M3 had very similar predictions for white applicants. In contrast, the expected likelihood of approval for minority applicants was quite different between M2 and M3. As mentioned above, in M2, the likelihood of approval for minority applicants was strongly associated with neighborhood racial composition. In M3, there was no statistical association between minority approval rates and neighborhood racial composition after the cross level interaction was included in the model (hence the straight light grey solid line in Figure 5.2).

The geographic contingent lending model (M3) showed that racial and socioeconomic aspects of neighborhoods were significant predictors of the impact of race of applicant. (In fact, the direct effect of the race of applicant was insignificant after neighborhood level variables were included.) As a neighborhood increases its socioeconomic status (50% to 80% white was the peak likelihood of approval), a white applicant is more likely to be approved. Including these variables increases overall model fit, but reduces the significance of the neighborhood racial composition variables impact on the level-1 intercept (π_{0n}). Thus, we can conclude that in this dataset, the importance of neighborhood racial composition was associated with the impact of the race of the applicant rather than for all applicants in general.

Table 5.4 Results of M2 and M3 for 2000-2002 Cohort

	Model 2		Model 3	
Fixed Effect	Coefficient	*p*-value	Coefficient	*p*-value
INTRCPT	0.860473	0.024	1.45725	<0.001
PW	0.023596	0.013	-	-
NS	0.069441	0.678	0.189306	0.001
NW2	-0.0002	0.003	-	-
NS2	-0.06733	<0.001	-0.05715	<0.001
NWNS	0.003641	0.06	-	-
OCC	0.714479	<0.001	0.709661	<0.001
FRS	0.740166	<0.001	0.733228	<0.001
FDIC	0.900777	<0.001	0.89962	<0.001
OTS	0.924813	<0.001	0.925442	<0.001
NCUA	0.571793	0.015	0.605235	0.01
LCON	-0.56034	<0.001	-0.55059	<0.001
LINC	-0.3486	0.191	-0.42123	0.113
LLOAN	1.035841	<0.001	1.079124	<0.001
PNHW	0.398845	<0.001	-0.23037	0.371
PW	-	-	0.027314	<0.001
NS	-	-	0.192367	<0.001
NW2	-	-	-0.00023	<0.001
CNHW	0.062378	0.131	0.057867	0.162
PF	0.162003	<0.001	0.152869	<0.001
LTI	-3.2788	<0.001	-3.52013	<0.001
Y1	0.179803	<0.001	0.180417	<0.001
Y2	0.081363	0.042	0.081915	0.041
FV	-4.4614	<0.001	-4.73371	<0.001

Table 5.5 shows the percentage of level-2 variance explained for each of the three models tested: the comparison model (M1), the expanded model (M2), and the geographically contingent model (M3). It shows that each subsequent model accounted for slightly more of the level-2 variance and each slightly reduced the spatial correlation in the level-2 error terms.

Table 5.5 2000-2002 Model fit statistics

2000-2002 Loan Cohort	Model		
	M1	M2	M3
Percentage of Level-2 Variance Explained	75.7%	76.0%	76.5%
Spatial Correlation	0.679	0.657	0.643
Change in Spatial Correlation from Unconditional Model	-0.134	-0.157	-0.170

Figure 5.2 Relationship between neighborhood racial composition and probability of approval for white and minority applicants for each of the 2000-2002 models

2003-2004 Loan Cohort – M2 – Expanded Model

Further exploration of the second loan cohort showed that vacancy rates and

population change were significant predictors of loan approval. Including these variables

slightly reduced the coefficient on neighborhood racial composition and its significance

level. Thus, there was weaker evidence for redlining with this specification. There was also no evidence of an interaction between neighborhood SES and neighborhood racial composition. However, at level-1, applicant race was still an important predictor of loan approval. The best model can be found in the table below. Looking at Figure 5.3, we see a simple downward shift from the black lines (M1) to the dark grey lines (M2), but no discernible change in impact of neighborhood percent white.

2003-2004 Loan Cohort – M3 – Geographic Contingent Lending

In M3, as with the earlier 2000-2002 cohort, both neighborhood racial composition and socioeconomic status were significant predictors of the importance of the race of applicant and when these interaction effects were included, the direct impact of applicant race was insignificant. White applicants were more likely to be approved in whiter and wealthier neighborhoods although this effect was non-linear (the negative coefficients on the squared terms showing that this was ameliorated slightly when in the most white and wealthy neighborhoods). This final model did not increase the amount of level-2 variance explained in the model, but substantially reduced the spatial correlation of the level-2 error terms (see Table 5.7). Figure 5.3 shows, unlike in the 2000-2002 data, that the predicted likelihood of approval for white applicants in M3 had quite a different trajectory from M2. There was a sharp drop in approval probability for white households applying in high minority neighborhoods (10% white or less), with approval peaking between 40% and 80% white before dropping down again to almost being equal with minority applicants in 100% white neighborhoods. Once again, minority applicants weren't favored in comparison to whites in any neighborhoods but for those that were almost 0% white.

111

Table 5.6 Results of M2 and M3 for 2003-2004 Cohort

	Model 2		Model 3	
Fixed Effect	Coefficient	p-value	Coefficient	p-value
INTRCPT	1.295356	<0.001	1.428466	<0.001
PW	0.011509	0.055	-	-
VAC	-0.0312	0.026	-0.03593	0.009
POPC	0.292092	0.024	0.276824	0.024
NS	0.164917	0.002	-0.07105	0.075
NW2	-0.00013	0.011	-	-
NS2	-0.02497	0.085	-	-
OCC	0.394436	<0.001	0.389804	<0.001
FRS	0.777583	<0.001	0.759082	<0.001
FDIC	0.950897	<0.001	0.930894	<0.001
OTS	0.20635	0.089	0.203773	0.094
NCUA	0.517107	0.058	0.601448	0.028
LCON	-0.31679	<0.001	-0.31376	<0.001
LINC	-0.52513	0.047	-0.62309	0.018
LLOAN	1.012066	<0.001	1.051772	<0.001
PNHW	0.365729	<0.001	-0.4016	0.19
PW	-	-	0.034402	<0.001
NS	-	-	0.329383	<0.001
NW2	-	-	-0.0003	<0.001
NS2	-	-	-0.04011	0.008
CNHW	0.073448	0.145	0.069924	0.166
PF	0.031961	0.516	0.01396	0.776
LTI	-3.35593	<0.001	-3.6208	<0.001
FV	-5.55731	<0.001	-6.44689	<0.001
Y1	0.084383	0.043	0.077874	0.062

Table 5.7 2003-2004 Model fit statistics

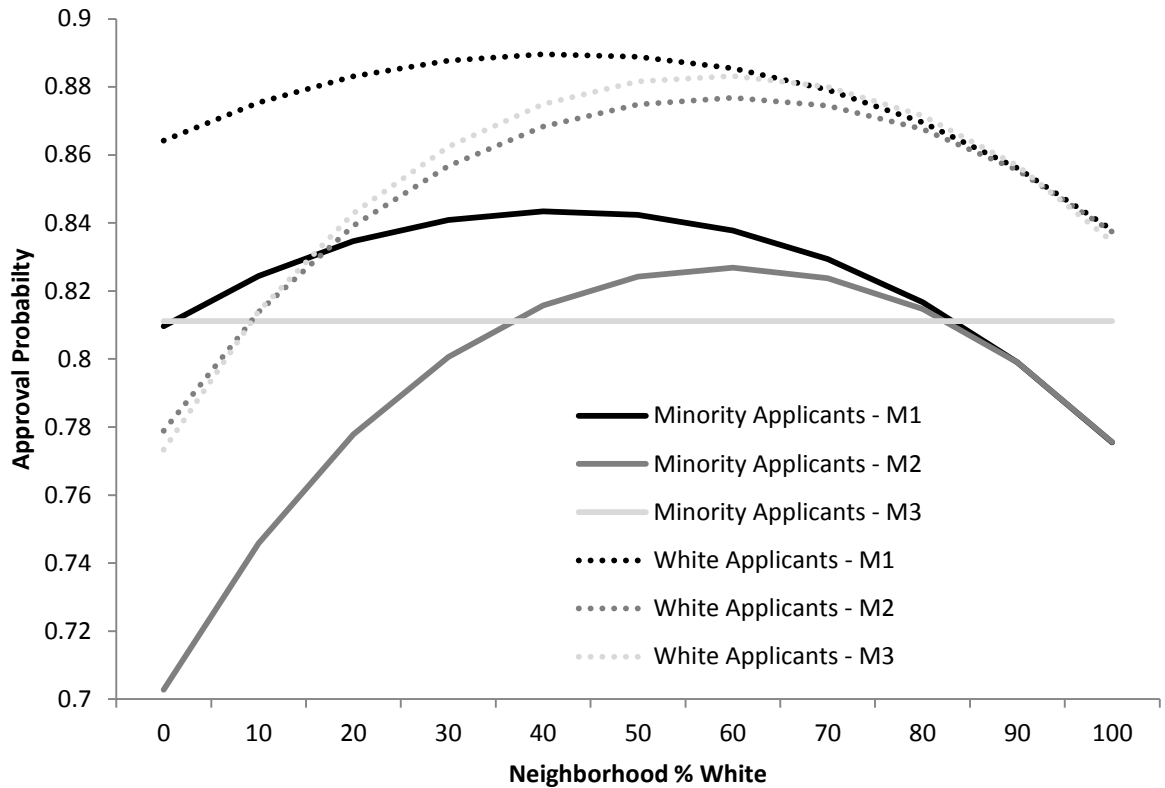| 2003-2004 Loan Cohort	Model		
	M1	M2	M3
Percentage of Level 2 Variance Explained	67.5%	67.8%	67.1%
Spatial Correlation	0.524	0.489	0.256
Change in Spatial Correlation from Unconditional Model	-0.357	-0.392	-0.625

Figure 5.3 Relationship between neighborhood racial composition and probability of approval for white and minority applicants for each of the 2003-2004 models

2005-2009 Loan Cohort – M2 – Expanded Model

Further analysis showed that there was a significant interaction between neighborhood racial composition and neighborhood socioeconomic status. Population change was just short of traditional significance thresholds and vacancy rates did not appear to have any relationship with the dependent variable. This fuller model continued to show evidence of neighborhood redlining and discrimination on the basis of the race of applicant (H1 and H2). As can be seen in Table 5.9, this model accounted for 5% more of the level-2 variance, but it increased the spatial dependence of the model. Figure 5.4

shows how the importance of neighborhood racial composition increases after the interaction term and population change variable were included. The approval probability for minority applicants in M2 increases from below 0.6 in a 0% white neighborhood to over 0.75 in a 100% white neighborhood (it peaks at 60-70% white).

2005-2009 Loan Cohort – M3 – Geographic Contingent Model

In M3 for the 2005-2009 loan cohort, unlike the earlier two analyses, neighborhood racial composition did not moderate the impact of race of applicant on approval probability. Thus in Figure 5.4 the light and dark grey lines run right over one another. However, neighborhood socioeconomic status was positively associated with approval probability. In this analysis, white applicants fared better in all neighborhoods. Also, in contrast to the earlier cohorts, there was evidence of redlining in the 2005-2009 data with the racial composition variables continuing to predict the level-1 intercept coefficient. Compared to M2, this model accounted for slightly less of the level-2 variation, but reduced the spatial dependence to a larger extent.

Table 5.8 Results of M2 and M3 for 2005-2009 Cohort

Fixed Effect	Model 2 Coefficient	Model 2 *p*-value	Model 3 Coefficient	Model 3 *p*-value
INTRCPT	0.383066	0.205	0.37675	0.211
PW	0.034458	<0.001	0.034982	<0.001
POPC	0.212158	0.061	0.203796	0.072
NS	-0.17368	0.194	-0.20404	0.128
PW2	-0.00027	<0.001	-0.00027	<0.001
NS2	-0.03601	0.002	-0.03443	0.003
NWNS	0.004759	0.007	0.004132	0.019
OCC	0.156741	<0.001	0.156986	<0.001
FRS	0.347993	<0.001	0.341001	<0.001
FDIC	0.937528	<0.001	0.92965	<0.001
OTS	0.199585	0.009	0.203269	0.008
NCUA	0.10779	0.489	0.15346	0.327
LCON	-0.29368	<0.001	-0.28992	<0.001
LINC	-1.0501	<0.001	-1.08208	<0.001
LLOAN	1.265015	<0.001	1.27646	<0.001
PNHW	0.471303	<0.001	0.476093	<0.001
NS	-	-	0.098123	<0.001
CNHW	0.230444	<0.001	0.224266	<0.001
PF	0.083876	0.011	0.078111	0.019
LTI	-4.93914	<0.001	-5.02855	<0.001
Y1	-0.11621	<0.001	-0.11526	<0.001
Y2	0.007564	0.415	0.008464	0.362
Y3	0.032358	<0.001	0.032438	<0.001
FV	-3.08265	0.008	-3.63526	0.002

Table 5.9 2005-2009 Model fit statistics

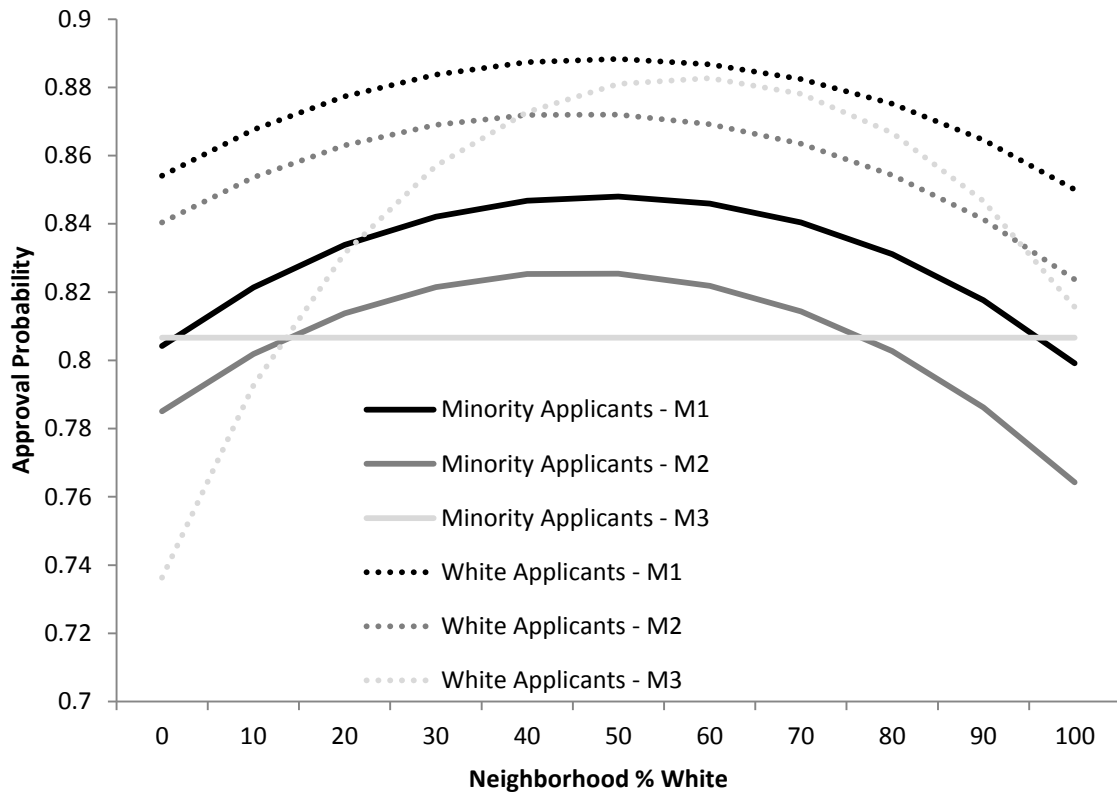| 2005-2009 Loan Cohort	Model M1	Model M2	Model M3
Percentage of Level 2 Variance Explained	42.4%	47.8%	47.4%
Spatial Correlation	0.323	0.364	0.346
Change in Spatial Correlation from Unconditional Model	-0.549	-0.508	-0.526
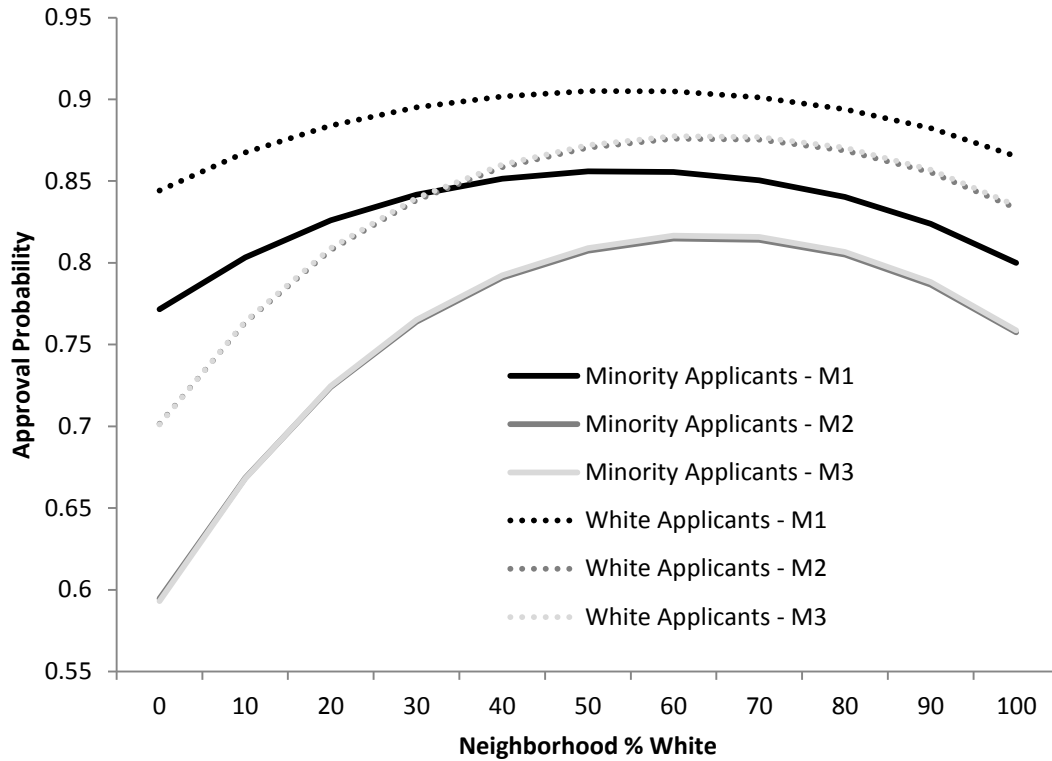
Figure 5.4 Relationship between neighborhood racial composition and probability of approval for white and minority applicants for each of the 2005-2009 models

Discussion

The analysis in Chapter 4 attempted to quantify the tract level impact of lending decisions on neighborhood racial composition; the current analysis attempted to quantify the impact of applicant race and neighborhood racial composition on the likelihood of loan approval. The results show that white applicants are more likely to be approved in almost all neighborhoods and that neighborhood racial composition matters, but its impact is not strictly linear.

The study looked at three loan cohorts and each was examined in three ways. The first (M1) examined all three cohorts with the same set of variables. The second (M2) introduced more neighborhood level variables, and removed any level-2 variables that were not strongly associated with the likelihood of approval. The third (M3) examined whether neighborhood level variables (neighborhood racial composition and neighborhood SES) predicted the impact of applicant race on the likelihood of approval (testing for geographic contingent lending).

While there was some variation in results from loan cohort to loan cohort, the predictions of the models generally stayed similar when the relationships between the key variables were graphed. White applicants were almost always more likely to be approved, and loans in neighborhoods with mixed racial compositions (approximately 50%-80% white) had the highest likelihood of approval.

The results of the neighborhood contingent lending model (M3) for the two earlier cohorts (2000-2002, 2003-2004) were similar to those of Holloway's research. That is, lenders did not seem to be making lending decisions based on the race of applicant nor the racial composition of the neighborhood alone, but on the interaction between them. The third cohort (2005-2009) did not show evidence of neighborhood racial composition impacting the likelihood of approval based on race. In this cohort, instead, neighborhood racial composition was consistently important for all applicants (although increasing neighborhood SES did improve the likelihood of white applicant approval).

While there was a clear importance of neighborhood racial composition on the likelihood of approval for many of the models, it is not clear that this phenomenon would

117

be best described as "redlining". In all models the impact of neighborhood racial composition was non-linear and the likelihood of loan approval tended to peak in neighborhoods with mixed racial compositions. In some of the models the whitest neighborhoods had about the same likelihood of approval as the least white neighborhoods. This perhaps would be predicted by research by Wyly and Hammel (1999) that showed that mortgage capital was flooding back to innercity neighborhoods. In Nashville, lenders do not seem to shy away from investing in non-all white neighborhoods, this could be due to the growing tendency of innercity revitalization and gentrification within Nashville over the past decade.

While there does seem to be a statistical bias in favor of white applicants and mixed neighborhoods, the limitations of HMDA data means these findings need to be interpreted with some caution. Lending decisions are based on more information than is available in these analyses, and it could be that confounding variables (e.g., credit score) that are not in this data set could account for some of the differences between the approval decisions of white and minority applicants. However, since these data are not made available by lenders - and that the analysis of the one dataset that included this information (the Boston Fed Study) found biases in favor of whites (Munnell, et al., 1996) – it seems that the burden of proof belongs to those who believe that there is no systematic racial bias in lending.

CHAPTER VI

DISCUSSION AND CONCLUSION

This dissertation began by introducing the three most widely cited explanations of contemporary racial residential segregation. It then proposed three research questions that attempted to explore the adequacy of these explanations. Before exploring the details of each question, a close examination of preference-based models of segregation was provided in Chapter 2. Chapter 3 described an empirical examination of whether homebuyer behavior in the Nashville metropolitan area was in line with what would be predicted by preference theorists. The second analysis, examined in Chapter 4, was similar to the first, but instead of looking at whether homebuyer locational decisions would lead to high levels of segregation, it focused on the lending decisions of banks and whether these were having an impact on racial residential segregation. The third analysis, which used individual loan applications rather than the aggregated characteristics of applicants by year and tract, tested whether lenders seemed to have a pattern of discrimination based on the race of applicant, the neighborhood racial composition, or the interaction between the two. The method and results of this analysis were described in Chapter 5.

From a purely theoretical standpoint, preference-based theories provide the most coherent explanation of racial residential segregation. The theoretical models that started with Schelling (1971), and developed further by Fossett (2006), provide a powerful

explanatory tool for why cities may have the high levels of segregation they do, even if the majority of households may prefer to live in relatively mixed neighborhoods. Fossett's sophisticated computer program demonstrates that segregation would be the "natural" outcome of housing decisions under plausible and empirically based assumptions about neighborhood preferences. The thought-provoking results of his and Schelling's theoretical models are described in detail in Chapter 2.

While preference-based theories of residential segregation have strong theoretical underpinnings, their empirical validation has been elusive. Chapter 3 described the results of a study that aims to fill this void by examining whether actual homebuyer decisions are in line with what would be predicted by preference-based models of segregation. The study, using a spatial hierarchical linear model, examined the percentage of home loan applicants who were white in each census tract in each year between 2000 and 2010, and compared this to the neighborhood racial composition of the census tract they are applying in.

Chapter 3 showed that an increasing proportion of applicants were white as the neighborhood percentage of whites increased, which was in line with preference-based models. However, the overall results suggested that the pattern of home loan applications would be more likely to lead to greater integration than greater segregation. This finding becomes clear when the predicted proportion of applicants who were white was examined by neighborhood racial composition.

The general pattern of homebuyers examined suggested that mixed and minority neighborhoods (a census tract with less than 80% white) would become whiter, while the

whitest neighborhoods are predicted to have in-movers that were less white than the existing neighborhood racial composition. If preference-based theories of "neighborhood tipping" were correct, one would expect minority and mixed neighborhoods to have a profile of applicants less white than those currently residing in the neighborhood, and the whitest neighborhoods to have a profile of applicants who were as white or whiter than those already residing in the neighborhood.

The findings of Chapter 3 suggested that that the theories of Schelling and Fossett do not provide an adequate account of the behavior of homebuyers in the Nashville metropolitan area during the 2000s. The findings also showed little support for socioeconomic based explanations of segregation. While median applicant income was positively associated with percentage of applicants who were white, this effect was relatively small. Further, neighborhood socio-economic status was negatively associated with the percentage of applicants who were white. SES-based explanations would predict that as neighborhood status increased so would the percentage of applicants who were white. This suggests that the socioeconomic-based explanations, as recently emphasized by Brown and Chung (Brown & Chung, 2006; Brown & Chung, 2008; Chung & Brown, 2007), still fail to account for many aspects of locational decisions by race.

It is important to note that Chapter 3 does not attempt to model the effect of racial biases in the housing market that may increase segregation. However, the limited support for preference and socio-economic based explanations suggests that it would be unwise to discount the importance of institutional racial biases in the housing market playing a role in maintaining contemporary segregation. These conclusions echo those of Freeman (2000) and Adelman (2005), who find that preference and socioeconomic factors are

121

unable to explain the segregation of minorities in the Multi-City Study of Urban Inequality data. The following chapter, Chapter 4, explicitly looked into the role of lenders in maintaining segregation.

The analysis in Chapter 4 explored the influence of lenders on racial residential segregation by looking at whether neighborhood racial composition was associated with lending decisions making a pool of applicants more or less white. For example, if lenders approved a disproportionate share of white applicants in white neighborhoods it would suggest a role of lending institutions in maintaining segregation.

The results provided a complex picture of the relationship between lending decisions and neighborhood racial composition. While lending decisions did not have a consistent impact on tracts across time, there was a statistical association between neighborhood percent white and the extent that lenders favored/disfavored white applicants. When all control variables were included at their means for each neighborhood composition, all neighborhoods saw a positive differential in white successful applicants. Neighborhoods that were between 30% and 50% white had the strongest impact on the differential approval rate of white applicants. When the impact on the actual number of white households moving into a neighborhood was taken into account, neighborhoods that had white populations between 60% and 90% saw the greatest increase in white households. The impact of these lending decisions is the equivalent of making a census tract six one-hundredths of a percent whiter.

These findings suggest a minor, if complex, relationship between lending decisions and segregation. They do not, however, provide evidence of racial bias (it

could be that the lending decisions are based on widely accepted underwriting criteria). Chapter 5 examines the data more closely to see whether there is evidence congruent with discrimination that may lead to segregation.

Unlike the first two analyses, the third used individual loan applications as the basic unit of analysis. It examined whether there is association between applicant race and neighborhood racial composition with the likelihood of loan approval, while controlling for a number of lender, loan, applicant, and neighborhood characteristics. Previous research has suggested that lenders may be more likely to deny applications from minority households who are moving into white neighborhoods: a pattern that is in line with the results described in Chapter 4.

Chapter 5 tested three different racial bias hypotheses: discrimination based on the race of applicant, discrimination based on neighborhood racial composition (redlining), and discrimination based on the interaction between race of applicant and neighborhood racial composition (geographic contingent lending). Due to the size of the dataset, the analysis was run in three separate cohorts. In the first two cohorts, including data from 2000 through 2004, the results suggested that white applicants were favored in all neighborhoods but for those that had almost no white residents. White applicants were most likely to be approved in neighborhoods that had between 50% and 80% of their residents white. The third cohort of loans, including applications made between 2005 and 2009, showed evidence of redlining and individual discrimination but not geographic contingent lending.

While there are some distinctions between the results found in the three different loan cohorts examined in Chapter 5, graphing the results showed that the predictive models offer substantially similar conclusions in most cases. White applicants in neighborhoods that have residents that are between 50% and 80% white are most likely to be approved. Like the findings of the Boston Fed Study (Munnell, et al., 1996), white applicants are favored over minority applicants in nearly all circumstances. It is clear that neighborhood racial composition and applicant race seem to have an impact on the likelihood of a loan application approval. However, unlike the findings of Holloway (1998) and Holloway and Wyly (2001), it is also clear that applying in the whitest neighborhoods does not increase the likelihood of approval (compared to mixed neighborhoods). This could be due to the general higher costs associated with many of these neighborhoods, making it more difficult to qualify for a loan.

The limitation of analysis 3 is that the full information available to lenders is not made public through HMDA data, which means that the findings are merely suggestive of biases. However, given the consistent findings of bias in favor of white applicants, lenders would be wise to provide more information to show that there were legitimate financial reasons for a higher rate of approval for white households.

In summary, the combination of the analyses presented in Chapters 3, 4 and 5 indicate that: preference theories do not provide an adequate explanation of the aggregate behavior of homebuyers in Nashville; white applicants are more likely to have loans approved when compared to non-whites; neighborhood racial composition has an impact on lending decisions, but it is not linear.

These results provide a strong empirical critique of Schelling and Fossett's preference models of segregation. They also provide little support for socioeconomic factors being the driver of racial residential segregation. Thus, the ultimate conclusion of these analyses suggests that the role of housing institutions in maintaining segregation should continue to be examined and extended into a stronger theoretical framework. Preference theories have gained their prominence through a highly developed theoretical analysis; in contrast, explanations focusing on racial biases tend to have an abundance of empirical support but lack a cohesive theoretical framework to tie disparate findings together. An inductively built theory, based on findings like those in Chapters 4 and 5 and the findings of the most recent national Housing Discrimination Study will help policy makers understand the relationship between a diffuse pattern of racial biases and widespread segregation (much like preference theories are able to tie isolated individual decisions to high levels of segregation). Similarly, a stronger theoretical framework connecting institutional biases to segregation would help counteract the intuitive appeal of socioeconomic explanations that have continued prominence despite very little empirical support.

As noted in Chapter 1, there is evidence of racial bias at every stage of the housing search. This study looks at one particular stage towards the very end of homebuyers' search for housing. It could be the case that the differential marketing of homes based on neighborhood racial composition (M. Collins & Galster, 1995) combined with the increase in racial steering by real estate agents (Galster & Godfrey, 2005) and the differential treatment by lenders before loan applications are made (Ross, et al., 2008), could have shaped the racial composition of applicants before the data used in this

study is even collected. Further research that examines the actual impact of these patterns on the distribution of households by race would buttress and expand upon this present study and help provide an empirical framework for theorizing the driving factors of contemporary racial residential segregation.

One other important finding that was consistent throughout all three analyses was the significance of the neighborhood racial composition-neighborhood SES interaction control variable. Higher SES neighborhoods tend to accentuate the importance of neighborhood racial composition on each of the dependent variables under study. While the weight of the findings within this study do not show a lot of support for direct socioeconomic explanations of racial residential segregation, it is clear that socioeconomic neighborhood factors do play a role in racial patterning. Further explorations of the connection between how neighborhood socioeconomic factors interact with neighborhood racial composition in determining household and institutional decision making as it relates to segregation would be a highly productive area of further study. These studies may also provide an opportunity to specifically test the importance of SES-based explanations of segregation, something that was beyond the scope of this study.

At a less grand scale, empirical research on the locational patterns of renters by race and the patterns of out-movers (rather than those moving into a neighborhood) would help build on the findings of Chapter 3. Closer examination using both quantitative and qualitative research methods on the patterns of lending approvals in neighborhoods that show consistent patterns of lenders favoring particular racial groups would help flesh out the complicated findings presented in Chapters 4 (and, to a lesser extent, Chapter 5).

While there seems to be a systematic relationship between neighborhood racial

composition and the decision making of lenders, traditional notions of redlining, where

banks withhold credit from mixed and non-white neighborhoods, seem to be out of date.

Fleshing out the larger consequences of these new lending patterns for urban

development and the geography of opportunity is an ongoing research project[19] that

would help scholars and policy makers develop better approaches to ensuring the original

goals of fair lending legislation like the Community Reinvestment Act and urban

revitalization programs like HOPE VI and Choice Neighborhoods are used in the most

productive way.

---

[19] Elvin Wyly is a leader in this area

APPENDIX

AN EXAMPLE OF CALCULATING CENSUS 2000 CHARACTERISTICS

FOR 1990 CENSUS GEOGRAPHIES

If 1990 tracts "A" and "B" split into three tracts in 2000: "X", "Y" and "Z". And, all of

X's population lived in the geography covered by A, and all of Y's population was

contained with B, but Z was split 60/40 between A and B.

- To calculate A's population characteristics in 2000:
    - I added X's characteristics to 60% of Z's. If both X and Z had a white
      population of 100 each in 2000, the newly calculated "A in 2000" would
      have a white population of 160 (i.e., all of X's and 60% of Z's white
      population).
- To calculate median income and median home value
    - I multiplied the median value by the percentage of the census tracts
      population that made up the original 1990 tract. For example, if tract X
      (median income = $40,000) made up 75% of A and tract Z (median
      income = $50,000) made up the remaining 25%, then I multiplied each by
      its proportion of the 1990 tract and then added those two numbers
      together. In this case it would be $40,000 x .75 (i.e., $30,000) + $50,000 x
      .25 (i.e., $12,500) = $42,500.

REFERENCES

Abariotes, A., Ahuja, S., Feldman, H., Johnson, C., Subaiya, L., Tiller, N., et al. (1993). *Disparities in Mortgage Lending in the Upper Midwest: Summary of Results Using 1992 Home Mortgage Disclosure Act Data*. Paper presented at the Fannie Mae University Colloquium on Race, Poverty, and Housing Policy.

Acevedo-Garcia, D., Lochner, K. A., Osypuk, T. L., & Subramanian, S. V. (2003). Future Directions in Residential Segregation and Health Research: A Multilevel Approach. *American Journal of Public Health, 93*(2), 215-221.

Adelman, R. M. (2005). The Roles of Race, Class, and Residential Preferences in the Neighborhood Racial Composition of Middle-Class Blacks and Whites. *Social Science Quarterly, 86*(1), 209-228.

Alba, R. D., & Logan, J. R. (1992). Assimilation and Stratification in the Homeownership Patterns of Racial and Ethnic Groups. *International Migration Review, 26*(4), 1314-1341.

Alba, R. D., Logan, J. R., & Stults, B. J. (2000). How Segregated Are Middle-Class African Americans? *Social Problems, 47*(4), 543-558.

Avery, R. B., Brevoort, K. P., & Canner, G. B. (2007). Opportunities and issues in using HMDA data. *Journal of Real Estate Research, 29*(4), 351-379.

Brown, L. A., & Chung, S.-Y. (2006). Spatial segregation, segregation indices and the geographical perspective. *Population, Space and Place, 12*(2), 125-143.

Brown, L. A., & Chung, S.-Y. (2008). Market-Led Pluralism: Rethinking Our Understanding of Racial/Ethnic Spatial Patterning in U.S. Cities. *Annals of the Association of American Geographers, 98*(1), 180-212.

Bullard, R. D. (Ed.). (2007). *The black metropolis in the twenty-first century*. New York: Rowman & Littlefield.

Card, D., Mas, A., & Rothstein, J. (2007). Tipping and the Dynamics of Segregation: National Bureau of Economic Research, Inc.

Charles, C. Z. (2003). The dynamics of racial residential segregation. *Annual Review of Sociology, 29*(1), 167-207.

Chung, S.-Y., & Brown, L. (2007). Racial/Ethnic Residential Sorting in Spatial Context: Testing the Explanatory Frameworks1. *Urban Geography, 28*(4), 312-339.

Clark, W. A. V. (2002). Ethnic Preferences and Ethnic Perceptions in Multi-Ethnic Settings. *Urban Geography, 23*(3), 237-256.

Clark, W. A. V. (2006). Ethnic Preferences and Residential Segregation: A Commentary on Outcomes from Agent-Based Modeling. *Journal of Mathematical Sociology, 30*(3 & 4), 319 - 326.

Clark, W. A. V., & Fossett, M. (2008). Understanding the social context of the Schelling segregation model. *PNAS, 105*(11), 4109-4114.

Collins, C. A., & Williams, D. R. (1999). Segregation and Mortality: The Deadly Effects of Racism? *Sociological Forum, 14*(3), 495-523.

Collins, M., & Galster, G. (1995). Discriminatory marketing of racially mixed neighborhoods in Memphis. *Journal of Urban Affairs, 17*(4), 339-355.

Crowder, K., South, S., & Chavez, E. (2006). Wealth, Race, and Inter-Neighborhood Migration. *American Sociological Review, 71*(1), 72.

Cutler, D. M., & Glaeser, E. L. (1997). Are Ghettos Good or Bad? *The Quarterly Journal of Economics, 112*(3), 827-872.

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (2008). When are ghettos bad? Lessons from immigrant segregation in the United States. *Journal of Urban Economics, 63*(3), 759-774.

Dawkins, C. J. (2004). Recent Evidence on the Continuing Causes of Black-White Residential Segregation.

Dawkins, C. J., Shen, Q., & Sanchez, T. W. (2005). Race, space, and unemployment duration. *Journal of Urban Economics, 58*(1), 91-113.

Echenique, F., & Fryer, R. G. (2007). A Measure of Segregation Based on Social Interactions. *Quarterly Journal of Economics, 122*(2), 441-485.

Fischer, M. J., & Massey, D. S. (2004). The Ecology of Racial Discrimination. *City and Community, 3*(3), 221-241.

Flippen, C. (2004). Unequal Returns to Housing Investments? A Study of Real Housing Appreciation among Black, White, and Hispanic Households. *Social Forces, 82*(4), 1523-1551.

Fossett, M. (2006a). Ethnic preferences, social distance dynamics, and residential segregation: theoretical explorations using simulation analysis. *Journal of Mathematical Sociology, 30*(3).

Fossett, M. (2006b). Including Preference and Social Distance Dynamics in Multi-Factor Theories of Segregation. *The Journal of Mathematical Sociology, 30*(3 & 4), 289 - 298.

Fossett, M. (2011). Generative Models of Segregation: Investigating Model-Generated Patterns of Residential Segregation by Ethnicity and Socioeconomic Status. *The Journal of Mathematical Sociology, 35*(1-3), 114.

Fossett, M., & Dietrich, D. R. (2009). Effects of city size, shape, and form and neighborhood size and shape in agent-based models of residential segregation: are Schelling-style preference effects robust? *Environment & Planning B, 36*, 149-169.

Fossett, M., & Waren, W. (2005). Overlooked Implications of Ethnic Preferences for Residential Segregation in Agent-based Models. *Urban Studies, 42*(11), 1893.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2000). *Quantitative Geography*. London: Sage.

Freeman, L. (2000). Minority Housing Segregation: A Test of Three Perspectives. *Journal of Urban Affairs, 22*(1), 15-35.

Galster, G., & Godfrey, E. (2005). By words and deeds - Racial steering by real estate agents in the US in 2000. *Journal of the American Planning Association, 71*(3), 251-268.

Goering, J. (2006). Shelling redux: How sociology fails to make progress in building and empirically testing complex causal models regarding race and residence. *Journal of Mathematical Sociology, 30*(3), 299 - 317.

Grady, S. C. (2006). Racial disparities in low birthweight and the contribution of residential segregation: A multilevel analysis. *Social Science & Medicine, 63*(12), 3013-3029.

Grady, S. C., & McLafferty, S. (2007). Segregation, Nativity, and health: Reproductive health inequalities for immigrant and native-born black women in New York City. *Urban Geography, 24*(4), 377-397.

Holloway, S. R. (1998). Exploring the Neighborhood Contingency of Race Discrimination in Mortgage Lending in Columbus, Ohio. *Annals of the Association of American Geographers, 88*(2), 252-276.

Holloway, S. R., & Wyly, E. K. (2001). "The Color of Money" Expanded: Geographically Contingent Mortgage Lending in Atlanta. *Journal of Housing Research, 12*(1), 55-90.

Jones, R. (2008). The Ambiguous Roles of Suburbanization and Immigration in Ethnic Segregation: the Case of San Antonio. *Urban Geography, 29*(3), 196-223.

Kim, S. (2000). Race and home price appreciation in urban neighborhoods: evidence from milwaukee,wisconsin. *Review of Black Political Economy, 28*(2), 9.

Kim, S. (2003). Long-Term Appreciation of Owner-Occupied Single-Family House Prices in Milwaukee Neighborhoods. *Urban Geography, 24*(3), 212-231.

Krivo, L. J., & Kaufman, R. L. (1999). How Low Can It Go? Declining Black-White Segregation in a Multiethnic Context. *Demography, 36*(1), 93-109.

Krivo, L. J., & Kaufman, R. L. (2004). Housing and Wealth Inequality: Racial-Ethnic Differences in Home Equity in the United States. *Demography, 41*(3), 585-605.

Laurie, A. J., & Jaggi, N. K. (2003). Role of 'vision' in neighbourhood racial segregation: A variant of the Schelling segregation model. *Urban Studies, 40*(13), 2687-2704.

Logan, J. R. (2001). *Ethnic Diversity Grows, Neighborhood Integration Lags Behind*. Albany: Lewis Mumford Center for Comparitive Urban and Regional Research.

Logan, J. R., Alba, R. D., & Leung, S. Y. (1996). Minority Access to White Suburbs: A Multiregional Comparison. *Social Forces, 74*(3), 851-881.

Macpherson, & Sirmans. (2001). Neighborhood Diversity and House-Price Appreciation. *The Journal of Real Estate Finance and Economics, 22*(1), 81-97.

Macy, M., & Rijt, A. v. d. (2006). Ethnic preferences and residential segregation: theoretical explorations beyond Detroit. *Journal of Mathematical Sociology, 30*(3), 275-288.

Massey, D. S., & Lundy, G. (2001). Use of Black English and Racial Discrimination in Urban Housing Markets: New Methods and Findings. *Urban Affairs Review, 36*(4), 452.

Munnell, A. H., Geoffrey, M. B. T., Lynn, E. B., & James, M. (1996). Mortgage Lending in Boston: Interpreting HMDA Data. *American Economic Review, 86*(1), 25-53.

Pattillo, M. (2005). Black middle class neighborhoods. *Annual Review of Sociology, 31*, 305-329.

Pulido, L. (2000). Rethinking Environmental Racism: White Privilege and Urban Development in Southern California. *Annals of the Association of American Geographers, 90*(1), 12-40.

Quercia, R. G., McCarthy, G. W., Ryznar, R. M., & Talen, A. C. (2000). Spatio-Temporal Measurement of House Price Appreciation in Underserved Areas. *Journal of Housing Research, 11*(1), 1-28.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierachical linear models: Applications and data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Rosenbaum, E., & Friedman, S. (2007). *The housing divide: how generations of immigrants fare in New York's housing market*. New York: New York University Press.

Ross, S. L., Turner, M. A., Godfrey, E., & Smith, R. R. (2008). Mortgage lending in Chicago and Los Angeles: A paired testing study of the pre-application process. *Journal of Urban Economics, 63*(3), 902-919.

Ross, S. L., & Yinger, J. (2002). *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair Lending Enforcement*: MIT Press.

Savitz, N. V., & Raudenbush, S. W. (2009). Exploiting spatial dependence to improve measurement of neighborhood social processes. *Sociological Methodology, 39*(1), 151-183.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*(1), 143-186.

South, S. J., Crowder, K., & Chavez, E. (2005). Geographic mobility and spatial assimilation among U. S. latino immigrants. *The International migration review, 39*(3), 577-607.

South, S. J., & Crowder, K. D. (1998). Leaving the'Hood: Residential Mobility between Black, White, and Integrated Neighborhoods. *American Sociological Review, 63*(1), 17-26.

Squires, G. D., & Kubrin, C. E. (2006). *Privileged places : race, residence, and the structure of opportunity*. Boulder, Colo.: Lynne Rienner Publishers.

St. John, C. (2006). Comment on Mark Fossett's Manuscript: Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations Using Simulation Analysis. *The Journal of Mathematical Sociology, 30*(3 & 4), 327 - 332.

Tootell, G. M. B. (1996). Redlining in Boston: do mortgage lenders discriminate against neighborhoods? (Vol. 96). Boston: Federal Reserve Bank of Boston.

Turner, M. A., & Ross, S. L. (2005). How Racial Discrimination Affects the Searchfor Housing. *The Geography Of Opportunity: Race And Housing Choice In Metropolitan America.*

Turner, M. A., Ross, S. L., Galster, G. C., & Yinger, J. (2002). Discrimination in Metropolitan Housing Markets. *Report to the US Department of Housing and urban Development. Washington, DC: Urban Institute.*

Wasserman, H., & Yohe, G. (2001). Segregation and the provision of spatially defined local public goods. *American Economist, 45*(2), 13-24.

Wyly, E. K., & Hammel, D. J. (1999). Islands of Decay in Seas of Renewal: Housing Policy and the Resurgence of Gentrification. *Housing Policy Debate, 10*(4), 711-771.

Zhang, J. (2004). A dynamic model of residential segregation. *The Journal of Mathematical Sociology, 28*(3), 147-170.