

UCLA

UCLA Electronic Theses and Dissertations

Title

Matching Residential Mobility to Raced-Based Neighborhood Preferences in Los Angeles:
Implications for Racial Residential Segregation

Permalink

<https://escholarship.org/uc/item/9174793s>

Author

Jarvis, Benjamin Fletcher

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Matching Residential Mobility to
Raced-Based Neighborhood Preferences in Los Angeles:
Implications for Racial Residential Segregation

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Sociology

by

Benjamin Fletcher Jarvis

2015

© Copyright by
Benjamin Fletcher Jarvis
2015

ABSTRACT OF THE DISSERTATION

Matching Residential Mobility to Raced-Based Neighborhood Preferences in Los Angeles: Implications for Racial Residential Segregation

by

Benjamin Fletcher Jarvis

Doctor of Philosophy in Sociology

University of California, Los Angeles, 2015

Professor Jennie E. Brand, Chair

Neighborhood racial composition preferences have the potential to produce extreme residential segregation between racial groups. Whites' preferences to avoid Blacks, and Blacks' preferences to maintain residential contact with their own group can induce high levels of White-Black segregation. Similar preferences can lead to segregation for Asians and Latinos. These theoretical facts have lead some observers to attribute continued residential segregation in the United States to preferences. These observers imply that elimination of housing market discrimination, socioeconomic differences, and other discrepancies between groups in metropolitan housing markets, would do little to lessen residential segregation by race. Most support for this perspective 1) relies on caricatures of empirically observed preferences, or 2) fails to examine whether preferences align with actual housing choices. These limitations are problematic. A failure to include realistic assessments of preferences in simulation models can lead to an overstatement of the potential for preferences to generate segregation, and a corresponding understatement of the potential for housing market discrimination and

other race-related factors to induce and sustain segregation. This dissertation addresses the limitations of previous attempts to vet the importance of preferences for segregation. It develops an appropriate statistical model for comparing stated racial composition preferences to actual neighborhood attainments. It uses this model to assess whether people migrate in ways that agree with their preferences, or whether some groups are systematically frustrated in matching their neighborhood attainments to their preferences. Finally, it combines results from these empirical assessments in simulation models of segregation. When some groups are stymied in migrating according to their stated preferences, does this tend to result in higher levels of segregation? Using data from Los Angeles County, this dissertation shows that Latinos, and to a lesser degree, Blacks, are disadvantaged relative to Whites in matching their residential preferences to their attainments. In simulation, these frustrated preferences yield excess levels of racial segregation: If people were able to migrate according to their stated preferences, levels of residential segregation would be approximately 20% lower. These results illustrate that preferences set a floor for levels of segregation, but do not fully explain extant segregation. Room remains for explanations based on discrimination and social networks.

The dissertation of Benjamin Fletcher Jarvis is approved.

Mark S. Handcock

Ka-Yuet Liu

Jennie E. Brand, Committee Chair

University of California, Los Angeles

2015

For my father

I had to code a little bit, but not too much

Contents

Acknowledgements	xiv
Biographical Sketch	xxiii
1 Introduction	1
2 Wouldn't It Be Nice? Comparing Racial Composition Preferences to Residential Histories in Los Angeles	14
2.1 Theoretical Perspectives on Segregation	16
2.1.1 Mismatches between Neighborhood Preferences and Attainments . . .	18
2.2 Theoretical Expectations	22
2.3 Data: Stated Preferences and Residential History in Los Angeles County . .	23
2.3.1 Stated Preference Data in L.A.FANS	24
2.3.2 RH: Combining Individual-Level Data in L.A.FANS with Neighborhood-Level Data from the Census	26
2.3.3 RH: Individual Level Controls	30
2.4 A Discrete Choice Approach to combining SP and RH Neighborhood Choice Data	32
2.5 Model Results	39
2.5.1 Predicted Probabilities and Wald Tests, by Respondent Race	39
2.5.2 Latino Respondents	40

2.5.3	White Respondents	42
2.5.4	Black Respondents	45
2.5.5	The Effect of Non-Racial Controls	48
2.5.6	Putting RH-SP Mismatches in Context	48
2.6	Discussion and Conclusion	52
3	Discrete Choice Models for the Joint Analysis of Residential Preference and Residential History Data	57
3.1	Data for Assessing Racial Composition Influences on Neighborhood Choices: Stated Preferences and Residential History	59
3.1.1	Stated Preference Data	59
3.1.2	Residential History Data	60
3.1.3	Contrasting Strengths and Weaknesses of SP and RH Data	62
3.2	Combining SP and RH Data: A Critical Review	63
3.2.1	Previous Work Combining SP and RH Data	64
3.2.2	The Importance of Choice Sets	65
3.2.3	Multi-Dimensionality in Neighborhood Choice	67
3.2.4	Micro-Macro Linkages	69
3.3	Discrete Choice Models for Combining SP and RH Data	69
3.3.1	Discrete Choice: Conditional Logistic Regression	70
3.3.2	Conditional Logit Models for Combining SP and RH Data	73
3.3.3	Combining Data from Multiple SP Sources	76
3.3.4	Accounting for Multiple Observations within Respondents: Mixed Logit	77
3.4	Examples: Whites in L.A.FANS	82
3.4.1	Models of Rank SP Data	82
3.4.2	Joint Models of Rank SP and Ideal SP Data	85
3.4.3	Joint Models of RH and Rank SP	88
3.5	Conclusion	91

4	Do Stated Preferences and Residential Histories Imply the Same Levels of Segregation? Testing for Divergence Using a Schelling Model	93
4.1	Background: Schelling Models and the Importance of Preferences	96
4.2	Methods: Schelling Model Specification	99
4.2.1	Structural Specification	99
4.2.2	Behavioral Specification	100
4.2.3	Measuring Segregation	105
4.3	Results	107
4.3.1	Attaining Equilibrium	108
4.3.2	Qualitative View of Segregation	109
4.3.3	SP-RH Differences: Entropy	110
4.3.4	Decomposing the Entropy Index	111
4.3.5	Dissimilarity Indices	113
4.4	Discussion and Conclusion	115
5	Conclusion	119
	Tables	129
	Figures	150
	Bibliography	182

List of Tables

2.1	Descriptives Statistics for Los Angeles Neighborhoods, 1998 and 2008	129
2.2	Descriptive Statistics for L.A.FANS Respondents and Their Families	130
2.3	Descriptive Statistics for L.A.FANS Respondents' Stated Neighborhood Pref- erences	131
2.4	Descriptive Statistics for L.A.FANS Respondents' RH Neighborhoods	132
2.5	Racial Composition Coefficients from Joint SP and RH Conditional Logistic Regression Models of Neighborhood Choice	133
2.6	Socioeconomic, Demographic, and Proximity Coefficients from Joint SP and RH Conditional Logistic Regression Models of Neighborhood Choice	135
2.7	Wald Tests of Racial Composition Effects Within RH and SP Neighborhood Choice Scenarios	137
2.8	Wald Tests of Differences in Racial Composition Effects: SP vs. RH Neigh- borhood Choice Scenarios	138
2.9	Cross Scenario Dissimilarities in Predicted Probabilities from Joint SP-RH Models of Neighborhood Choice, Los Angeles Neighborhood Choice Set . . .	139
2.10	Predicted Neighborhood Racial Composition based on a Los Angeles Neigh- borhood Choice Set, from Conditional Logit Models of SP and RH Data . .	140
3.1	Racial Composition Coefficients from Conditional Logit Models of Whites' Rank SP	141

3.2	Racial Composition Coefficients from Conditional Logistic and Mixed Logit Models of Whites' Rank SP	142
3.3	Estimated Covariances of Racial Composition Coefficients for Whites, Mixed Logit Models of Rank SP Data	143
3.4	Racial Composition Coefficients from Conditional and Mixed Logit Models of Whites' Neighborhood Choices in Rank SP and Ideal SP Vignettes	144
3.5	Estimated Covariances of Racial Composition Coefficients for Whites, Mixed Logit Models of Rank and Ideal SP Data	145
3.6	Racial Composition Coefficients from Conditional and Mixed Logit Models of Whites' Neighborhood Choices in RH and Rank SP Vignettes	146
3.7	Estimated Covariances of Racial Composition Coefficients for Whites, Mixed Logit Models of Rank SP and RH Data	147
4.1	Entropy Index by Scale and Source of Racial Composition Coefficients, from Schelling Segregation Models	148
4.2	Pairwise Indices of Dissimilarity by Scale and Source of Racial Composition Coefficients, Schelling Residential Segregation Models	149

List of Figures

2.1	Sample Ranked Neighborhood Vignette Card	150
2.2	Ideal Neighborhood Vignette Card	151
2.3	Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino	152
2.4	Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black	153
2.5	Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White	154
2.6	Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian	155
2.7	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino	156
2.8	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black	157
2.9	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White	158
2.10	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian	159
2.11	Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino	160

2.12	Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black	161
2.13	Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White	162
2.14	Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian	163
3.1	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian from Conditional Logit Models of Ranked SP Data	164
3.2	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black from Conditional Logit Models of Ranked SP Data	165
3.3	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino from Conditional Logit Models of Ranked SP Data	166
3.4	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White from Conditional Logit Models of Ranked SP Data	167
3.5	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian in Mixed Logit Models of Ranked SP Data	168
3.6	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black, Mixed Logit Models of Ranked SP Data	169
3.7	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino, Mixed Logit Models of Ranked SP Data	170
3.8	White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White, Mixed Logit Models of Ranked SP Data	171
4.1	Typical Grid Set-Up for Schelling Models	172
4.2	Neighborhood Definitions for Schelling Models	173
4.3	Initialization of City and Population Structure	174
4.4	Agent Behavioral Specification	175

4.5	Theil's Entropy Index vs. Simulation Time, by Source of Racial Composition	
	Coefficients and Scale Parameter	176
4.6	Dissimilarity Indices vs. Simulation Time, by Source of Racial Composition	
	Coefficients and Scale Parameter	177
4.7	RH Coefficient Based Simulations	178
4.8	Rank SP Coefficient Based Simulations	179
4.9	Theil's Entropy Index vs. Scale Parameter (θ), by Source of Racial Composi-	
	tion Coefficients	180
4.10	Dissimilarity Indices vs. Scale Parameter (θ), by Source of Racial Composition	
	Coefficients	181

Acknowledgements

Arriving here at the end of the dissertation, I am finding it almost as difficult to assemble my acknowledgements. So many people have contributed, both consciously and unwittingly, to this modest intellectual effort. I worry about leaving somebody out, or not giving enough credit where credit is due. And will my language be adequate to the task? The words that follow (or, rather, have come before) this preface tend toward the analytical, staid, and, at times, formulaic. But I feel I should make words of acknowledgement dance, lithe and light, as those I mention here deserve. Given the current state of my writing practice, I will likely fall short of this goal. But I think the people I mention here, and even those I leave out, are generous enough of heart to accept any shortcomings in my praise and prose.

I first must acknowledge the mentorship, advice, and thoughtful consideration that Rob Mare has given me along the way. Sudden illness prevented him from officially serving as committee chair, but he served as chair in everything but name, even dating back to before my admittance to UCLA. In 2008 I was possessed by the crazy idea that I should apply to graduate programs in the social sciences, born of the faint realization that I had already been doing social science work in business, and a strong sense that I was not working on the sort of social problems I thought were and are important. I contacted Rob because he seemed a well respected guy, as far as my uneducated eye could detect, and he was close by—I was already living in the Los Angeles area. Neophyte that I was, I had little notion of the respect he commanded among his colleagues, the depths of his contributions to the discipline, or the value of his time. Despite surely having more important things to do, he

responded to an initial message I sent him and agreed to meet me in person. He advised me on which programs fit with my nascent research interests, and gave me some insights into which scholars would best nurture my scholarship. I came away from the conversation sure that I would apply to UCLA, and if I were admitted, that I would work with Rob. Since then, Rob has given me the space to develop my own ideas, and a strong enough guiding hand to keep me from running off track. He has been very patient with me, tolerating my sometimes agenda-less and meandering meetings, and my frequently rambling prose. I have felt all along that I have had his confidence, attention, and backing, and that has meant a lot. He deserves a huge share of the credit for my development as a sociologist and I will be forever grateful.

I am immensely grateful to Jennie Brand for stepping in as the chair of the committee in Rob's absence. I do not mean to imply that her contributions were limited to this brief and pressing call of duty. Jennie has been there for me nearly as long as Rob has, dating back to a conversation we had before I matriculated. Since then, she has served as a reader for my master's thesis, recruited me into a project using data in the Census Research Data Center at UCLA, and provided me with important quantitative training. Jennie has a keen eye for shaping an analytical narrative, one I hope I can continue to benefit from as I work to revise these chapters into publications.

Mark Handcock is a busy man with many graduate students in statistics demanding his attention. Nonetheless, he found time for me, a sociology student with few formal ties to UCLA's Department of Statistics. It began with his network analysis course, which came as a revelation. Since then, he has opened my eyes to methodological approaches that stay true to the social processes they claim to represent. I have likely disappointed him as a statistician, and perhaps I remain too reluctant to enter the statistical weeds, but I benefited greatly from his teaching and his advice.

Ka-Yuet Liu did not have a lot of time to influence my work or my training, but I am thankful that she joined the department and agreed to join up with my committee. She

offered me several long conversations that have pushed me to think carefully about the simulations in this dissertation, and will certainly help me to develop the current batch of simulations into publishable work.

While on the subject of my committee, I should acknowledge the contributions of its secret fifth member, Judy Seltzer. While I never took a class with Judy, and my work often strayed from her family demography wheelhouse, Judy offered me numerous insights as I developed my research. From pressing me to think more carefully about preferences, how they are measured and how they correspond to behavior, to critiquing my writing, Judy has always provided me with thoughtful comments and a patient ear.

My peers in UCLA's Stratification, Demography, and Family Research Group have also aided me as well. This group, under the guidance of Rob Mare and Judy Seltzer, continues to serve as a forum in which graduate students present work in progress. While Rob and Judy provided many helpful comments during our meetings, I also benefited from the presence of my peers. They provided valuable insights and perhaps more valuable examples of how to conduct and present research. From the start, Charles Lau, one of the chief organizers, welcomed me into the fold, and then immediately set to work coaxing a National Science Foundation Graduate Research Fellowship proposal out of my disorganized research ideas. William Rosales also provided key comments on this early proposal and encouraged me to dive head first into research. The research group also provided opportunities for collaboration. In this regard, I formed a bond with Xi Song. While we have yet to make good on our project investigating patterns of intra-generational occupational mobility, I foresee useful research emerging from this collaboration soon. In the meantime, Xi has commented on my papers and presentations, offered spirited challenges to my weaker ideas, and pushed me to become a more concise and focused researcher. During weekly meetings of the research group, I have also received helpful comments from many other student participants, including Jenjira Yahirun, Yana Kucheva, Sung Park, Erika Arenas, Kate Choi, Claudia Solari, Rahim Kurwa, Sung Park, Bongoh Kye, and John Sullivan.

Many other faculty and staff, both within and outside UCLA, aided me as I developed the ideas, cleaned the data, and refined the methods contained in this dissertation. Cameron Campbell provided encouragement early on. Thanks to him, I also branched out to develop new projects, applying my methods to historical data on migration in the nineteenth century United States. I am also grateful to Patrick Heuveline for his approach to teaching, and for providing me with my early training in demographic methods. I also remember helpful conversations I had with him about simulation models. Anne Pebley helped me develop a better understanding of Los Angeles, the setting for this dissertation. I benefited from her expertise with the Los Angeles Family and Neighborhood Survey, and also from her knowledge of LA's neighborhoods and its social demography. In finding answers to my questions about the Los Angeles Family and Neighborhood Survey, Anne often passed me to Christine Peterson at RAND. Christine was always quick to respond to my e-mails, and found clear answers to my esoteric questions about the data.

Meghan Sweeney and Steve Clayman provided valuable comments on my early thesis ideas. Ying Nian Wu provided me with foundational statistical training. His lectures were cogent and captivating, and I still turn to notes from those lectures for occasional help in clarifying fundamental statistical concepts. Li Cai's course on statistical programming helped to solidify my technical and intuitive understanding of more advanced statistical models. I am also grateful for Mike Seltzer's course on causal inference, and his willingness to provide advice about my research projects and my graduate curriculum. Elizabeth Bruch's Stata code helped orient me with the Los Angeles Family and Neighborhood Survey. She also willingly ceded some research ground, which allowed me to pursue the research contained in these chapters. Wendy Fujinami helped me to navigate the UCLA bureaucracy, and ably swung into action when Rob fell ill. I still owe her Diddy Reese cookies. Nathan Osgood, from the University of Saskatchewan, provided a useful and informative workshop on agent based modeling at UCLA. He was crucial to getting the ball rolling on the simulation models presented here, and even provided answers to a few technical questions long after his duties

to his workshop participants had passed. Finally, I would like to thank David Grusky for providing me with a quiet space to work at Stanford as I tapped out drafts of my dissertation chapters.

I received support and advice from several other graduate student peers at UCLA. I am thankful for my cohort-mate and good friend Brooks Ambrose. I still miss the short lived, quasi-café that we ran out of his office before he moved to the San Francisco Bay Area. It was a great place to get a good cup of coffee (a difficult find on campus!), relax, gossip, and engage in some light theorizing. Across our many discussions he helped me hone my sociological thinking and challenged me to expand beyond my occasionally cynical and instrumentalist view of the social world. Together we also had a good, although not always productive, time pursuing our shared hobby: academic software. Amy Zhou was frequently there, enjoying the coffee with us. Amy Zhou has a keen mind for the study of race and ethnicity and helped me to vet some of the more antiquated race and ethnicity literature on my reading docket. She also encouraged me to push through difficult periods of writing, while providing a willing partner for the occasional and much needed run to acquire donuts to accompany our coffee. Yelizavetta Kofman would also occasionally join our salon. Liz has a way of scoffing at obstacles and making challenges seem eminently surmountable. Her attitude, I think, rubbed off on me just enough to get me through the dissertation process. I would also like to thank Juli Simon Thomas. We entered UCLA in the same cohort, and shared an interest in advanced quantitative methods. Together, we navigated the first three years of course work, helping each other to find appropriate classes, collaborating on problem sets, and preparing for our field exams. Pat Reilly pushed me to think more carefully about the cultural aspects of neighborhoods and migration, and his ethnographic sensibilities played a necessary counterpoint to the reductive tendencies of my demographic training. Space and time prevent me from expanding on their contributions, but I would also like to acknowledge Dwight Davis, Chris Rea, Winston Chou, Noah Grand, Krishna Bhogaonker, and Abigail Cooke.

Applying to graduate school was no easy feat for a person who had been out of school for six years. I would have been lost without the help of a few business colleagues and an old family friend. Vadim Korsunsky showed me the ropes in retail, taught me more than a few quantitative tricks, and helped me figure out how to present esoteric research to a wider audience. He held me to a high standard at work, taught me the importance of double and triple checking my tables, and was happy to provide a glowing recommendation when I needed it. Nick Burrows provided more tutelage in the retail arts, and taught me a great deal about how to prioritize outstanding issues at work. I hope to emulate his level headedness as I move forward in my career. Amy Farnham Walton was my first boss in industry, and I appreciated her humor and critical eye for detail. It speaks of her character that she was happy to write me a graduate school recommendation despite my having moved on to work for another company. Finally, Michael Kass, a near and dear family friend, looked over my graduate school application materials and gave me crucial advice about how to improve them. Michael has, in many ways, played the role of godfather since the death of my dad. He has expressed continuous interest in my whereabouts and goings on, and I have felt him rooting for me.

Speaking of family and friends, I have had more than my fair share willing to offer support. Of course we all owe the world to our mothers. My mom, Katherine Jarvis, provided all the love and support I could possibly have wanted throughout my elementary, secondary, and college years. Naturally she did not relent as I entered graduate school. She read drafts of my master's thesis and this dissertation, and I have her to thank for consistency in phrasing and relative rarity of typos, misspellings, and grammatical missteps in the final document. And of course she provided me with countless meals and a happy, warm place to come home to during holidays and the occasional weekend. Perhaps I am not outwardly grateful enough for her concern and pride, but my gratitude is great. I love you Mom.

I wish my dad, Pitts Jarvis, had lived to see me complete my doctorate. Despite his passing, I felt his presence as I worked, perhaps most resoundingly in my relative novice

status as a computer coder. I sometimes regret not having developed better programmer's chops, but I have benefitted immeasurably from his insistence that I nurture other skills. He did not want me merely to follow in his foot steps as a programmer and computer scientist, but rather to branch out into the humanities as well, becoming a critical social thinker in the process. I miss him dearly.

My sister Verity has been tremendously tolerant of my waywardness from the family home base in Palo Alto. She suffers from a bit of wanderlust, but has nonetheless stuck close to home to be with my mom as I have been in Los Angeles. I would be a bad, or at least sinfully unaware, family demographer if I did not acknowledge that this has been a sacrifice for her. But her sacrifice made me feel that I had the leeway to return to school.

Before graduate school, I was lucky enough to meet my partner and now wife Kiki Wang. It is at least partly, if not mostly, because of her (sorry, Rob) that I remained in Los Angeles. She has shown me so much love and patience as I have worked on the dissertation, and she has motivated me to persevere through the most stressful moments. Having her to come home to, and weekends with her to look forward to, were important parts of staying sane, happy, and productive. And now, thanks to this degree and Rob's looking out for his own, we now have new adventures to look forward to.

I should also acknowledge the backing and encouragement of my extended family. My grandmother, Suzanne Schall, and grandfather, Edmund Schall, kept me well fed in my undergraduate years, perhaps in anticipation of some lean years in graduate school. While my grandfather has passed away, and my grandmother may not have been cognizant of my graduate education, I am sure they both would have been proud to see me complete a degree that might land me in the academic fold, and potentially keep me closer to the orbit of an institution they both loved, Yale University. I would also like to thank my aunts, Vicki Escalera, Martha Czacskes, Laura Gouillart, and Margo Burwell, for providing material and emotional support. I hope I can continue to make them proud, and justify all the presents that they have sent my way over the years.

I can also thank many friends who provided recreational and emotional outlets, as well as the occasional meal. Salvador Delgado, no matter how harrowing his day job and the demands it placed on his time, found time to take me out to dinner and the occasional Dodgers' game. I did feel slightly guilty when his Dodgers would lose to my favored Giants, but I cannot reject the possibility that these very good years for Giants baseball rubbed off a bit on my own work. My friend Seth Johnson also provided key support. While his residence in Washington D.C. and my general reluctance to talk on the telephone with anyone limited our conversations to about once per year, it was always great to speak with him and use him as a sounding board. And visiting him was at least one motivation for attending my first Population Association of America conference in Washington D.C., even it was only to present a poster. Of his Mets, we need not speak.

Neil Van Os moved away from Los Angeles in the early years of my graduate education, but our get togethers and his enthusiasm for life and the outdoors is infectious. Thankfully he was not too far away, in Davis, CA, and Kiki and I were able to replenish our enthusiasm reserves by paying him visits. Conveniently enough, during these visits to Neil we were also able to see his partner, Jennifer Chen, whose wry wit, sharp intellect, and no nonsense attitude are inspirations in and of themselves. Through Neil, I met friend and fellow social scientist Robby Ratan. Robby was already well into graduate school when I began at UCLA. I have been lucky to receive his advice as I think strategically about my career. I would like to thank my long time friend Benn Myers, whose company was great comfort on my occasional forays back to Palo Alto for holidays and long weekends. I am also thankful for the support of Bret Ladine, Mansoor Razzaq, Daniel Lifshitz, David Rogier, Kieran Kennedy, Lauren Morel, and Paul Morel.

Lastly, I had financial backing from several sources. The Department of Sociology provided a backstop that I did not frequently have to turn to. UCLA's Grad Division provided ample support in the form of the Graduate Research Mentorship and Graduate Summer Research Mentorship programs, as well as the Dissertation Year Fellowship. I should also

acknowledge the California Center for Population Research at UCLA (CCPR) for its training support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (T32-HD007545).

Biographical Sketch

The author graduated from Palo Alto High School in 1999. He proceeded to Yale University where he completed a Bachelor of Science in Applied Physics in 2003. Thereafter, he joined Teach for America, serving as a ninth and eleventh grade physical science and physics teacher at Booker T. Washington High School in Atlanta, Georgia from 2003 until 2005. After his second year of teaching, he returned to his home state of California to work as a market researcher and marketing manager for McMaster-Carr Supply Company, an industrial supply company located in Santa Fe Springs, California. He spent two years working for the company, during which time he made a friend who later encouraged him to join up with a new grocery start up, Fresh and Easy Neighborhood Market, located in El Segundo, California. He worked as a consumer pricing analyst for Fresh and Easy until 2009, at which point he left to enter the doctoral program in sociology at the University of California, Los Angeles. In 2011 he was received a Master of Arts in Sociology. In 2015 he earned his Doctor of Philosophy in Sociology from the University of California, Los Angeles.

Chapter 1

Introduction

High degrees of residential segregation between racial and ethnic groups persist as defining features of the United States residential landscape (Iceland et al. 2013). Persistent residential segregation sparks concern because of its potential to affect, and in many cases restrict, the life chances of minority groups who are confined to racially homogeneous neighborhoods (Wilson 1987). This plays out in two ways. First, segregation consigns some racial and ethnic groups to neighborhoods of concentrated poverty and privileges others to concentrated affluence (Massey and Denton 1993; Massey and Fischer 2000; Quillian 2012). Second, those who live in minority neighborhoods may find their political clout constrained by segregated conditions. The deformation of political influence may inhibit efforts by residents of minority occupied neighborhoods to attract public and private resources and services (Logan 1978). In these ways, racial residential segregation may directly or indirectly influence a diverse set of outcomes, including educational attainment (Wodtke et al. 2011; Quillian 2014), family formation (Harding 2003), employment status (Wagmiller et al. 2006), mental health (Kling et al. 2007) and physical health (Collins and Williams 1999).

Factors Sustaining Segregation

Several factors have been tagged as likely culprits in sustaining segregation. Three sets of factors—socioeconomic differences, race-based residential preferences, and housing market discrimination—have received the bulk of the attention. Most studies suggest a limited or

inconsequential role for socioeconomic differences. This has left social scientists roughly split into two camps: those who advocate preference based explanations and those who advocate discrimination based explanations. However, a fourth set of factors, kinship and friendship networks, may also influence patterns of segregation. These factors have been given scant attention in the residential segregation literature. Here I review the theoretical understandings of how these factors can generate segregation, as well as evidence for their contributions to segregation.

Socioeconomic Differences

Socioeconomic stratification across racial groups can lead to racial segregation when neighborhoods have different levels of material advantages. People making housing choices in a competitive housing market will bid up housing prices in neighborhoods with widely desired amenities (e.g., well maintained housing stock, ready access to parks, or ocean views). Conversely, housing prices will plummet in neighborhoods with few amenities, or with undesirable characteristics (e.g., close proximity to polluting industrial plants, dilapidated housing, or limited access to retail). The resulting housing price stratification will tend to concentrate the poor, disproportionately drawn from racial and ethnic minority groups, into neighborhoods with few amenities. At the same time the rich, disproportionately drawn from more well to do groups, will concentrate into neighborhoods with the most desirable amenities. The three way interaction between spatially fixed amenities, economic disparities between racial groups, and competitive housing markets could thus induce racial segregation.

Evidence for the socioeconomic disparities explanation of segregation is thin. Income and education differences between Asians, Blacks, Latinos, and Whites account for some segregation between these groups, but not nearly all of it. Members of minorities with higher socioeconomic attainments do live in closer proximity to Whites. But even Blacks, Latinos, and Asians with high incomes are segregated from Whites (Denton and Massey 1988; Iceland and Wilkes 2006; Massey and Fischer 1999). And the benefits of socioeconomic attainment do

not accrue equally to groups. For example, Blacks with high income are less integrated with Whites than Latinos and Asians with low incomes (Denton and Massey 1988; Massey and Fischer 1999). Relatedly, income and education account for more White-Latino and White-Asian segregation than White-Black segregation (Bayer et al. 2004; Iceland and Wilkes 2006; Iceland et al. 2005; Massey and Fischer 1999). This evidence implies that other mechanisms work to sustain segregation—eliminating income and educational disparities would reduce segregation, but not eliminate it, especially for Blacks.

Housing Market Discrimination

Housing market discrimination, typically targeted at minority groups, could induce and sustain segregated outcomes, even with socioeconomic gaps closed. The mechanism is straightforward: if minorities wish to live in White neighborhoods, and can afford to do so, but are denied the opportunity to rent apartments or buy houses in those neighborhoods, then residential segregation by race is a natural consequence. The influence of discrimination remains something of an explanatory wild card. On one hand, high levels of racial segregation have persisted despite attempts to curtail institutional and market discrimination (e.g., the Fair Housing Act of 1968). On the other hand, discriminatory treatment is still exhibited by samples of housing market actors (Yinger 1995; Turner et al. 2002; Ross and Turner 2005; Turner et al. 2013). And even if audit methods currently used to assess housing market discrimination suggest recent declines in discrimination (Ross and Turner 2005), the locus of discrimination may be shifting to new domains as members of advantaged groups seek out and discover new legal ways to maintain their residential advantages (Massey 2005). However, it is unclear to what degree discriminating actors exert influence in the housing market. Housing seekers may bypass and marginalize the worst offenders such that these offenders ultimately affect few housing outcomes (Heckman 1998).

Assessing the degree with which housing market discrimination affects segregation is difficult because little or no data exist that tie together individual acts of discrimination

with inter-neighborhood migration events. Most assessments of the effect of discrimination on segregation have relied on metropolitan level data sets or data at the neighborhood level. Metro areas with higher incidence of discrimination, as measured by housing audit studies, also exhibit higher levels of Black-White segregation (Galster and Keeney 1988). And neighborhoods with higher levels of White representation are more likely to be sources of discrimination complaints (Galster 1987). These results suggest that discrimination deflects neighborhood choices for minorities, or at least it did during the 1970s and 1980s. However, using data on individuals reveals that residential mobility and neighborhood outcomes are not significantly associated with metropolitan level measures of discrimination (South and Crowder 1998a). In sum, studies that attempt to link segregation to discrimination have yielded mixed results.

Relying on aggregate measures of discrimination may be misleading because of potential measurement error and the ecological fallacy trap. Imputing discrimination at the metropolitan level to the experiences of individuals will create a noisy measure, and potential understatement of the effect of discrimination. But relying on metro-level assessments of discrimination could lead to a misstatement of their effects as well. For example, metropolitan areas with higher levels of housing discrimination may also contain Whites with more minority-averse residential preferences. These preferences, rather than housing market discrimination, could lead to higher levels of segregation.

Neighborhood Racial Composition Preferences

Given reductions in discrimination (Ross and Turner 2005) and the diversity of economic attainments within racial and ethnic groups, attention has increasingly focussed on the importance of preferences for patterning metropolitan levels of segregation (e.g., Clark 2009; Fossett 2006a; Clark and Fossett 2008). Between group differences in racial preferences can generate highly segregated residential patterns (Schelling 1969, 1971; Fossett 2006a). Segregation can emerge even if individuals tolerate non-trivial amounts of inter-racial mixing. In

other words, complex interactions between people can lead even relatively tolerant individuals to live in neighborhoods that are remarkably homogeneous in composition. Empirically observed racial residential preferences appear capable of generating and sustaining some degree of racial and ethnic segregation on their own (Bruch and Mare 2006; Clark 1991; Van De Rijt et al. 2009; Bruch and Mare 2009; Xie and Zhou 2012). Given this sufficiency, race-based preferences have garnered substantial attention, and they warrant more in depth consideration here.

Race-based residential preferences can stem from multiple sources. There are clear divisions along racial lines. Asians, Blacks, Latinos, and Whites on average tend to favor members of their own groups as neighbors relative to out-groups (Charles 2006, 2000). This has lead some to argue that benign in-group affinities, rooted in cultural and linguistic similarities, provide the basis for preferences (Clark 1991; Thernstrom and Thernstrom 1997). However, such benign preferences do not explain the tendency of all groups to adhere to a racial hierarchy when evaluating out-group neighbors—Whites are the most preferred out-group, followed by Asians, Latinos, and then Blacks (Charles 2000, 2006). Additionally, the variation in preferences within broad racial and ethnic categories, and the determinants of that variation, belie a simplistic cultural affinity explanation. Traditional demographic and socioeconomic variables, like age and education predict a small amount of variation in preferences. In addition, the degree to which individuals stereotype particular racial and ethnic groups and the neighborhoods in which they live also predicts variation in neighborhood racial composition preferences.

Preferences: Demographic and Socio-Economic Correlates

Traditional demographic and socioeconomic variables, like age, education, income, and nativity, are associated with racial preferences. Better educated Whites are more likely to express a willingness to integrate with Blacks, perhaps because those with more education are more likely to look beyond stereotypes and prejudices (Xie and Zhou 2012; Farley et al. 1997). Across multiple racial groups, those with more education are less desirous of own-race

neighbors, but also more likely to prefer White neighbors (Charles 2000). Younger Whites also tend to exhibit more tolerant preferences, perhaps because younger Whites have been raised in more racially and ethnically heterogeneous social environments than older cohorts, but also perhaps because of liberalizing trends in education over the course of the twentieth century (Quillian 1996). There is limited evidence that income and home ownership are associated with preferences net of other factors (Charles 2000). Finally, the foreign-born favor neighborhoods with more White residents and fewer Black residents, perhaps because they are more likely to give credence to an implicit racial hierarchy in the United States, in which Whites are perceived as the most prestigious group, and Blacks the least prestigious (Blumer 1958; Bobo 1999; Bobo and Johnson 2000). Overall, these demographic and socioeconomic variables appear to account for approximately half of the explained variation in residential preferences (Farley et al. 1994; Charles 2000).

Preferences and Stereotypes

Variation in negative stereotyping of racial and ethnic “others” also coincides with variation in neighborhood racial composition preferences. Individuals who negatively stereotype members of other racial and ethnic groups as unintelligent, welfare dependent, anti-social, discriminatory, and criminal are more averse to members of these groups as neighbors (Charles 2006, 2000). This is not to say that all preferences are rooted in prejudice. Providing some support for the benign in-group preferences hypothesis, those who express more positive feelings towards their own group have stronger preferences for own group neighbors relative to out-group neighbors (Krysan et al. 2009).

Rather than stereotyping individuals, people may stereotype the neighborhoods that certain groups occupy. For example, Whites use the presence of Blacks in a neighborhood to make inferences about levels of crime, degrees of physical deterioration, and housing price trajectories (Quillian and Pager 2001; Sampson and Raudenbush 2004; Ellen 2000; Krysan 2002; Harris 1999). It is unclear if these negative neighborhood stereotypes are easily separable from negative stereotypes of individuals (Krysan 2002; Krysan et al. 2009). More educated

individuals or those who are wary of being seen as bigoted may hide their negative stereotypes of Black *people* behind concern about housing prices or neighborhood deterioration in Black *neighborhoods*. In total, racial prejudices, stereotypes, and positive in-group feelings can account for the other half of the explained variation in preferences. However, much of the overall variance in racial preferences is unexplained by socio-demographic and social-psychological variables. Together, demographic, socioeconomic, and social-psychological variables account for less than one fifth of observed variation in preferences (Farley et al. 1994; Charles 2000).

Social Networks

Kinship, family ties, family structure, and social network based explanations of segregation have garnered relatively little attention compared to socioeconomic, discrimination, and preference based explanations. Their roles have received more attention in the literature on international migration and the development of ethnic enclaves (Boyd 1989; Choldin 1973; Massey et al. 1993; Massey and Espinosa 1997). The potential for kinship and family ties to perpetuate segregation is based on three social phenomena. First, people often migrate to places that are proximate to kin. The tendency to follow kin occurs at both the international and national scales of migration (Stark and Bloom 1985; Greenwood 1997; Massey et al. 1993), and it is natural to presume that some of the social network mechanisms that influence migration across long distances would also affect migration within cities. Within cities, the influence of social and kinship networks on neighborhood choices can take the form of co-residence: children moving in with parents, parents moving in with children, and friends or partners moving in with each other. Second, people frequently share the same racial and ethnic identifications as their family members, friends, and acquaintances. Third, residential segregation is already a fact of life in most, if not all, American cities. Combined, these three social phenomena can reproduce segregation. A tendency for people to maintain close spatial contact with family, friends and acquaintances, who frequently share the same racial identifications and who tend to live in neighborhoods that reflect the underlying segre-

gated metropolitan structure, will also sustain segregation, even if people eschew race-based residential preferences.

To some extent, social scientists are beginning to explore the potential for networks and family ties to shape intra-metropolitan migration. These studies build on previous research exploring the effect of chain migration, in which new migrants follow in the footsteps of previous migrants, on the formation of “ethnic neighborhoods” (MacDonald and MacDonald 1964). For example, Blacks, Latinos, and Whites appear to have different levels of knowledge about neighborhoods in Chicago (Krysan and Bader 2009). Awareness of communities is driven partly by the socioeconomic characteristics of individuals, but also by the racial composition of the communities in question, suggesting that racially patterned information networks may influence neighborhood awareness, and hence, migration. Meanwhile, household formation has the potential to influence levels of segregation. Mixed race couples appear to move to more diverse parts of cities, and in so doing may lower levels of between race segregation (Holloway et al. 2005; Ellis et al. 2012). The diversity seeking settlement patterns of mixed race couples ironically underscore the important role that racially homophilous dating and homogamy can play in sustaining segregation. Given the nascent character of this piece of the segregation literature, it is difficult to quantify exactly how patterns of social ties influence racial segregation. This is an area ripe for further research.

Evaluating the Contribution of Preferences to Segregation in Los Angeles

These four factors—socioeconomic differences, discrimination, preferences, and social networks—are likely not mutually exclusive determinants of segregation. Theoretically, they could, and likely do, work in concert to maintain levels of racial segregation, not to mention segregation along other (e.g., socioeconomic) lines. Indeed, the most complete theoretical models of segregation combine three of these factors—preferences, socioeconomic characteristics, and discrimination—in unified models (Clark and Fossett 2008; Fossett 2006a, 2011; Macy and

Van De Rijt 2006). These theoretical models find that socioeconomic sorting alone is not capable of producing high levels of racial segregation. Distributions of economic resources like income are not dissimilar enough across racial groups to generate the high degrees of segregation observed in many US metropolitan areas. In fact, in some simulated scenarios, socioeconomic sorting can lead to slightly less segregation, as people with low socioeconomic status are forced to live together, regardless of their group memberships or their preferences to avoid mixing. Instead, it appears that either between group preference differences or discrimination are necessary conditions for attaining high levels of segregation. Between group preference differences, on the whole, appear to be the more robust generators of segregation. In fact, under some conditions, discrimination alone is not even sufficient to generate high levels of segregation (Macy and Van De Rijt 2006). There are, however, reasons to believe that these theoretical models may be misleading. First, they frequently rely on out-dated caricatures of preferences when more realistic, up-to-date, empirically grounded preferences imply lower levels of segregation (Bruch and Mare 2006; Van De Rijt et al. 2009; Bruch and Mare 2009; Xie and Zhou 2012). This leaves space for other explanations of segregation to fill the theoretical void. Second, these studies have relied on quite limited specifications of discrimination and how it alters migration. It would be worth entertaining other specifications of discriminatory mechanisms. Third, these models have entirely ignored the role of social networks and kinship.

This dissertation addresses oversights in previous studies that connect segregation to between group preferences. It remains primarily concerned with between race differences in preferences for neighborhood racial composition, and their role in generating and sustaining high degrees of racial and ethnic residential segregation in the United States. However, unlike previous studies that port characterizations of race-based preferences into theoretical models of segregation (Fossett 2006a; Bruch and Mare 2006; Xie and Zhou 2012), I take the step of making an empirical comparison between race-based preferences and actual residential attainments. I determine whether race and racial composition seem to “matter” in the same

ways when people evaluate hypothetical neighborhoods, and when people migrate between neighborhoods in the real housing market. I then explore whether any differences I uncover imply different conclusions about the capacity of preferences to generate high degrees of segregation.

Comparing what people say they desire in a neighborhood to where they actually live acknowledges that discrimination and racially patterned social networks, frequently unobserved, can impose themselves between dreams and reality. Previous studies drawing the connection between preferences and segregation have paid insufficient attention to the possibility that preferences do not cleanly translate to the housing market. Some classes of people may face frustration matching their neighborhood attainments to their preferences. While previous simulation studies have incorporated non-racial factors, like housing quality and socioeconomic status, into their simulations, they also make implicit assumptions about the weight of preferences relative to unmeasured factors that are associated with race. Proponents of the preference based explanation of segregation are quick to point out that observed stated preferences are sufficient to sustain high levels of segregation, but they frequently fail to consider whether segregation would not, in fact, be higher if preferences were partially or wholly frustrated for some housing seekers (Fossett 2006a, 2011; Macy and Van De Rijt 2006).

As I have no data indicating how discrimination and social networks shape residential outcomes, I must remain agnostic as to which of these factors (if not others!) leads to mismatches between preferences and neighborhood attainments, and how these factors contribute to segregation net of preferences. This should not undermine the purpose of this dissertation. Even without explicitly accounting for these other factors, I can still carve out (additional) space for them in theories of segregation. Future studies can more carefully consider whether discrimination, social networks, or some other set of factors, occupies this space. With this in mind, the remaining chapters of this dissertation develop the argument as follows:

Chapter 2 performs the empirical comparison between preferences and neighborhood attainments. I use data from the Los Angeles Family and Neighborhood Survey to perform these comparison separately for four broad racial groups—Asian, Latino, White, and Black—occupying Los Angeles County. Several other scholars have explored the connection between race-based neighborhood preferences and neighborhood attainments (Ihlanfeldt and Scafidi 2002, 2004; Adelman 2005; Charles 2006; Clark 1992). My efforts diverge from this previous work in methods and substance. On the methods side, I assess mismatches between preferences and attainments using discrete choice models (Train 2009; Ben-Akiva and Lerman 1985). These models account for the fact that people face quite different neighborhood options when expressing preferences. The hypothetical neighborhoods people consider when stating preferences often do not correspond to many or any existing neighborhoods in the metro areas where they make real housing decisions.

In substance, my analysis explicitly addresses three confounding social processes that are only obliquely addressed, if at all, in previous studies. First, the analysis presented in this chapter recognizes the multi-racial reality of Los Angeles. Individuals and families negotiate a housing market in which neighborhoods are not just White or Black, but also feature Latinos and Asians with a variety of national origins and nativities. I consider the racial composition trade-offs people face across all of these groups simultaneously. Second, my analysis explicitly models the matching of people to neighborhoods based on socioeconomic status. In previous studies, only the demand side of this matching process, i.e., socioeconomic status at the individual or family level, is considered. The discrete choice models permit the inclusion of socioeconomic resources at both the individual and neighborhood level, and better represent the process of matching neighborhood supply to demand. Third, previous analyses have virtually ignored the spatial contingency of residential choices: when people move, they tend to move to spatially proximate neighborhoods. Given the spatial clustering of residential segregation, this will tend to retain people in disproportionately racially homogeneous parts of their cities. This suggests that the effect of race and racial composition on migration could

be a spurious result of pre-existing spatial patterns of segregation. My models account for this possibility.

This chapter finds that minority groups, namely Latinos and Blacks, are less successful than Whites in moving to neighborhoods that match their racial composition preferences. This finding is not distorted by the correlation between socioeconomic status and race, either at the neighborhood, individual, or family level. The conclusion is also robust to the inclusion of variables that account for the spatial contingency of segregation, and the tendency of people to migrate to spatially proximate neighborhoods.

Chapter 3 develops the methods that I use for the empirical comparison performed in Chapter 2. These methods are elaborations of discrete choice models originally developed in the econometrics literature (McFadden 1978). I adapt them to accommodate both preference data and real residential mobility data. In developing these models, I also present elaborations that can be used to characterize heterogeneity in preferences and residential mobility processes within racial groups. I present example applications of these models to data from the Los Angeles Family and Neighborhood Survey.

Chapter 4 takes the estimates from Chapter 2 and works out the implications of neighborhood preference-neighborhood attainment mismatches for patterns of segregation. This chapter relies on Schelling simulation models of segregation. The simulations presented in this chapter stipulate a population of actors who are divided into Latino, White, Black, and Asian groups, with proportions roughly equal to those observed in Los Angeles County. These agents populate a stylized, artificial city, which is composed of a grid of housing units. The agents migrate between housing units in the grid according to racial composition preferences that differ across the four groups, as observed in Chapter 2. I perform three sets of simulations. In the first two sets of simulations agents are assigned preferences based on two neighborhood vignette experiments contained in L.A.FANS and analyzed in Chapter 2. These simulations reflect the patterns of segregation that obtain when members of each group migrate according to their “pure” residential preferences. In the third simulation set,

agents move between neighborhoods based on the effect of racial composition on neighborhood choices observed in the residential histories of Los Angeles County residents. These simulations reflect not only preferences, but also the unobserved influence of racial discrimination, social networks, and other factors that impede people in moving to the neighborhoods they say they prefer. This chapter reveals that while pure preferences do yield high levels of segregation, levels of segregation are about 25% higher when agents are assigned preferences based on the actual influence of race and racial composition on migration in the housing market. This implies that there is still substantial space for explanations of segregation that are not wholly based on preferences. Finally, Chapter 5 synthesizes results from Chapters 2, 3, and 4, and suggests directions for future research.

Chapter 2

Wouldn't It Be Nice? Comparing Racial Composition Preferences to Residential Histories in Los Angeles

Relatively small differences between racial groups in their preferences for neighborhood level racial mixing can lead to complete residential segregation (Schelling 1971). Given the persistence of racial residential segregation in the United States despite the passage of the Fair Housing Act (Logan et al. 2004; Logan 2013; Massey and Denton 1993), this theoretical insight has animated substantial scholarly interest in race-based residential preferences. Direct assessments of preferences using survey-based instruments reveal preference differences between Blacks and Whites in the United States. Blacks prefer neighborhoods that are evenly split between Black and White residents; Whites express distaste for neighborhoods beyond 20% Black representation (e.g., Farley et al. 1978, 1993). These differences are consistent with some degree of segregation. However, the degree of segregation implied by preference differences across groups depends critically on their functional form and variations in these preferences within groups (Bruch and Mare 2006, 2009; Van De Rijt et al. 2009; Xie and Zhou 2012).

Preference-based explanations of segregation rest not only on assumptions about group differences in preferences and the shape of these preferences, but also on the assumption that housing seekers make residential choices in a free and fair housing market—members of each group move, without hindrance, to neighborhoods that match their preferences. But

migration is hardly free or unhindered. In addition to the direct financial and psychological costs of undertaking a move, people are subject to budget constraints when choosing new housing. And for some groups, having money and a will to move may not be enough. Empirical evidence from housing audit studies suggests that racial minorities, especially Blacks, face discrimination across metropolitan housing markets in the United States (Yinger 1995; Turner et al. 2002). Real estate agents, landlords, and other housing market actors provide minorities with inferior customer service, steer them away from White neighborhoods, withhold financial counseling, and appear to deny financing to otherwise qualified housing seekers (Reibel 2000; Williams et al. 2005). This lingering disadvantage renders it critically important to examine whether actual housing choices align with respondents' underlying racial composition preferences. Studies that use observed preferences, typically assessed using survey instruments, to chart out implications for aggregate patterns of segregation may be misstating the potential for segregation when ignoring mismatches between neighborhood preferences and actual neighborhood choices.

I investigate the correspondence between racial preferences and neighborhood attainments using data from the Los Angeles Family and Neighborhood Survey (L.A.FANS). This data set provides a unique opportunity to examine the relationship between neighborhood racial preferences and residential mobility because it contains two different kinds of data. First, L.A.FANS implemented a version of a neighborhood vignette experiment to obtain *stated preference (SP)* data describing how racial composition influences Los Angeles County residents' willingness to live in hypothetical neighborhoods. Second, L.A.FANS also obtained longitudinal *residential history (RH)* data describing where in Los Angeles County respondents have lived. The RH data in L.A.FANS are ideal for the purposes of this study because, unlike other data with an SP component (e.g., the Multi-City Study of Urban Inequality), these data identify migration events, not simply cross-sectional residential locations.

Using these data, I ask whether individuals' stated racial preferences align with the racial composition of their actual neighborhoods. Are some racial and ethnic groups better able

to match neighborhood attainments to preferences? To answer this question, I estimate discrete choice models that combine SP and RH data. These models offer a number of advantages over previous, ordinary least squares regression-based attempts to assess the relationship between preferences and prior residential attainments. First, they allow me to investigate how multiple racial composition components influence neighborhood choices, and to examine the influences of these racial composition components simultaneously within respondents. Second, discrete choice models accept non-racial controls at both the individual and neighborhood level. This allows me to consider a question only obliquely addressed in previous studies: Can we account for differences between preferences and residential history by accounting for the matching of individuals to neighborhoods according to socioeconomic characteristics, as well as the likely restriction of housing search to spatially proximate neighborhoods? Finally, the discrete choice approach allows me to model SP and RH outcomes simultaneously, avoiding the complication of assessing the potentially reciprocal relationship between preferences and prior residential history.

The remainder of this chapter discusses two theoretical perspectives that offer insight into how racial segregation emerged and persists in the United States. I propose minor modifications to these perspectives to generate hypotheses about for whom and in what directions SP and RH diverge. I then discuss the L.A.FANS data I use to test these hypotheses and provide a brief overview of the discrete choice methods I use to analyze these data. I then present results from discrete choice models of SP and RH neighborhood choices. Finally, I review my hypotheses in light of my results and conclude.

2.1 Theoretical Perspectives on Segregation

Two broad theoretical perspectives, spatial assimilation and place stratification, guide explanations of racial and ethnic segregation. The spatial assimilation perspective considers the residential integration of minorities as a corollary to their integration in other social

dimensions (Massey 1985).¹ As members of minority groups secure employment, ascend in socioeconomic status, and adapt linguistically and culturally, so too do they come to live in neighborhoods populated by the majority. Proponents of the theory are not always clear about the causal direction of effects—does residential integration lead to socioeconomic and cultural assimilation, or is assimilation along socioeconomic and cultural lines a necessary precursor to residential integration?—but the cross sectional implications are clear: Those who attain socioeconomic parity with the majority, achieve facility with the majority’s language, and master the majority’s cultural symbols and touchstones, will also live in close proximity to the majority. Controlling for social differences between groups, especially socioeconomic differences, should account for the residential segregation between those groups.

The second theoretical perspective, place stratification, focuses on political conflict over neighborhoods differentiated according to spatially situated goods and services. According to this perspective, residents of relatively advantaged neighborhoods literally and figuratively police boundaries between their neighborhoods and less advantaged parts of the city (Logan 1978). Coupled with conflict and group threat theories of race and ethnicity (Blalock 1967; Blumer 1958; Bobo 1999), the place stratification perspective suggests that majority group members, who tend to enjoy the most socioeconomic advantages and occupy the most advantaged neighborhoods in urban space, actively defend their neighborhoods against encroachment on the part of people perceived to be undesirable or threatening, especially members of minority groups. In theory, majority members mount this defense out of fear that an influx of minorities will either undermine the quality of the majority’s neighborhoods, threaten the majority’s political control over spatially distributed resources and prerogatives, or both. This perspective emphasizes that achieving cultural and economic parity with the majority cannot guarantee spatial parity for minorities. Through processes of housing discrimination and selective migration, members of the majority will exclude minority groups from majority occupied neighborhoods, and maintain political control over spatial advantages (see e.g.

¹Massey provides an overview of the spatial assimilation view, which has its roots in classic, Chicago School urban ecology. Also see Sampson (2011).

Massey and Denton 1993).

Both the spatial assimilation and place stratification perspective pay insufficient attention to the social fact of racial and ethnic residential preferences (Charles 2000; Farley et al. 1978, 1993, 1997). The spatial assimilation perspective largely ignores these preferences. It assumes that people strive to live in the best, materially endowed neighborhoods they can, with residential segregation along racial and ethnic lines manifesting as a side-effect of other, often socioeconomic, differences between groups. The place stratification perspective considers racial and ethnic preferences, but attributes them primarily to the majority. The majority acts on preferences for low out-group contact by building up discriminatory institutions, and, when those fail, moving away from neighborhoods when minorities begin to move in. The place stratification perspective implicitly denies that minority group members have their own preferences, or that these preferences matter much for extant patterns of segregation.

The social fact of racial and ethnic preferences can be incorporated into the above theories of segregation, but doing so requires a slight re-orientation. Rather than examining actual residential outcomes in isolation from racial and ethnic preferences, we should examine outcomes relative to those preferences. Do residential preferences and residential attainments match, and if not, for whom and why?

2.1.1 Mismatches between Neighborhood Preferences and Attainments

Modified spatial assimilation and place stratification perspectives would both suggest that majority group members, i.e., Whites, will be most likely to match their stated preference for racial mixing with their residential experiences. They would also both predict that disadvantaged minorities, especially Blacks and Latinos, will be least likely to achieve an SP-RH match.

Available evidence is ambiguous as to whether Whites achieve SP-RH matches more often than members of non-White groups. Whites with greater tolerance of Black neighbors

are more likely to live in neighborhoods with higher Black representation (Adelman 2005; Charles 2006; Ihlanfeldt and Scafidi 2004). Among Blacks, preferences are also associated with neighborhood outcomes: Blacks who have stronger preferences for living with Whites tend to live in neighborhoods with more Whites (Adelman 2005; Charles 2006; Freeman 2000; Ihlanfeldt and Scafidi 2002). In fact, the effect of preferences on exposure to Whites appears to be as strong or stronger for Blacks and other non-Whites as compared to Whites, with the caveat that the preference measures used in these studies, while similar, are not directly comparable across groups. Despite accounting for preferences, previous analyses have found substantial between-group differences in racial composition exposures. Whites tend to live in predominantly White neighborhoods regardless of their preferences, while Blacks, even those with the strongest preferences for living in White neighborhoods, live in largely Black neighborhoods (Adelman 2005). Similar results hold for other non-White groups (Charles 2006; Clark 1992; Bruch 2012).

The spatial assimilation and place stratification perspectives imply different sources for these mismatches. A modified spatial assimilation perspective would suggest that the mismatch between neighborhood preferences and residential circumstances is caused by group differences in socioeconomic standing and a corresponding correlation between racial composition and socioeconomic attributes at the neighborhood level. Members of economically disadvantaged groups, like Blacks and Latinos, may be unable to afford the neighborhoods that match their preferences for greater exposure to Whites, because those with socioeconomic advantages, especially Whites, bid up the cost of housing in these neighborhoods. If this is the case, then accounting for the sorting of people into neighborhoods according to socioeconomic factors should be sufficient to account for SP-RH discrepancies. To date, evidence for this perspective is not encouraging. Certainly preferences appear as only one factor among many influencing neighborhood racial composition exposures. Education, income, and wealth play important roles as well; across all groups, those with greater education, income, and wealth attainments tend to live in Whiter neighborhoods (Adelman 2005; Charles

2006; Freeman 2000; Ihlanfeldt and Scafidi 2002). But when accounting for both preferences and socioeconomic attainments, Whites still have greater neighborhood-level exposure to Whites, and Blacks greater exposures to Blacks.

However, people do not move fluidly between neighborhoods based on racial and socioeconomic considerations alone. On one hand, the act of moving incurs costs, both financial and psychological, that tend to fix individuals in place, at least in the short term. The immobile may find that their neighborhoods' racial and ethnic compositions change, pushing that composition out of line with stated preferences. Thus, we might account for SP-RH discrepancies by distinguishing between movers and stayers. On the other hand, even those who move may give more consideration to neighborhoods that will cause the least disruption to family, work, and other social life. As economic geographers have long noted, those who migrate longer distances often endure greater personal and social disruption (Sjaastad 1962; Greenwood 1985). When people do move they will likely find spatially proximate neighborhoods more appealing because migration to these neighborhoods incurs fewer psychological and social costs. Given the spatial clustering of segregation (Massey and Denton 1988; Reardon et al. 2008), people may thus opt for slightly less optimal neighborhoods, in terms of racial composition, because nearby destinations are less disruptive in other areas of life. Thus, accounting for the spatial proximity of destination neighborhoods among movers may also explain some of the discrepancy between preferences and the racial composition of destination neighborhoods. Scant research examines the relationship between individual residential mobility, neighborhood racial composition, and the spatial proximity of destinations. One study that considers all three reveals that Whites who live in neighborhoods and communities with high minority concentrations are more likely to engage in long distance moves (Crowder and South 2008). However, the OLS approach taken in this study, which treats distance as a dependent variable, does not consider the relative tradeoffs between migration distance on one hand, and the racial composition of destination neighborhoods on the other.

The place stratification perspective suggests that housing market discrimination may be a primary contributor to SP-RH discrepancies for some groups. Housing audit studies have uncovered disparate treatment by members of different racial groups at the hands of housing market actors like real estate agents (Yinger 1995; Turner et al. 2002). In general, members of racial minority groups receive worse treatment than Whites and are impeded in gaining access to White neighborhoods. This discrimination appears to extend deep into the real estate market, with mortgage brokers and underwriters also engaging in discriminatory practices targeted at racial and ethnic minorities (Reibel 2000; Williams et al. 2005). These discriminatory practices can directly impede minority groups in obtaining residence in their preferred neighborhoods. However, it is currently difficult, if not impossible, to directly test the effects of discrimination on SP-RH mismatches. Doing so would require joint observation of residential history, stated preferences, and instances of housing market discrimination. To my knowledge, there are no data sets that contain all three. In observational studies of residential mobility, one must deploy controls as best as one can, and then stake a claim to residual SP-RH differences.

There is an important flip side to the discrimination and socioeconomic disadvantage explanations of SP-RH mismatches: discrimination and socioeconomic disparities may per-versely contribute to SP-RH mismatches among advantaged as well as disadvantaged groups. First, institutional discrimination in rental and housing markets is carried out primarily by housing market actors like landlords, real estate agents, and mortgage brokers, not by the general population of advantaged housing seekers or by the settled population in any particular neighborhood. If housing market actors are more discriminatory, or less discriminatory, than the advantaged groups (i.e., Whites) on whose behalf they discriminate, then the resulting degree of neighborhood segregation may not reflect the preferences of the advantaged group. Second, if advantaged Whites (for example) have very strong preferences for segregation, but the housing market is in fact open, then the price of White neighborhoods may be bid up and predominantly White neighborhoods will only be accessible to those with

the highest incomes. Disadvantaged Whites may find themselves living in neighborhoods out of line with their preferences for segregation. Third, if advantaged Whites don't have preferences for segregation, but do have preferences for neighborhood amenities and features associated with socioeconomic characteristics, this, in combination with the relative economic advantages of Whites relative to many minority groups, may lead Whites to live in neighborhoods that are more White than they would like, but which better match their non-racial preferences. This means that, absent controls for matching of people to neighborhoods along socioeconomic lines, we might expect SP-RH discrepancies among Whites as well.

2.2 Theoretical Expectations

The above theoretical discussion need only be combined with some basic social facts about the relative statuses of racial groups in Los Angeles to derive a few testable hypotheses. First, Blacks and Latinos are the most socioeconomically disadvantaged populations in Los Angeles, while Whites and Asian groups are relatively economically advantaged (Treiman and Lee 1996). In addition, Blacks and Latinos face strong negative stereotypes (Bobo and Zubrinski 1996; Zubrinsky and Bobo 1996; Charles 2006), indicating that they will be most likely to face discrimination in their housing choices. This suggests the following hypotheses:

1. Blacks and Latinos will have the largest SP-RH discrepancies. In particular, they will have lower levels of co-residence with Whites compared to their preferences, and higher levels of exposure to disadvantaged groups, namely other Blacks and Latinos.
2. Whites' preferences will more closely match their neighborhood attainments than Blacks and Latinos. Whites will tend to prefer and live in neighborhoods with high proportions of White residents, and low proportions of Latinos and Blacks.
3. If individual and family-level socioeconomic factors prevent people from matching preferences to residential attainments, then controls for matching of individuals' economic circumstances to neighborhood socioeconomic circumstances should reduce or elimi-

nate the degree of SP-RH mismatch.

4. To the extent that migration incurs economic and social costs that people mitigate by either staying in place, or moving to spatially proximate neighborhoods, then distinguishing between movers and non-movers, as well as controlling for the spatial proximity of destination neighborhoods, should account for RH-SP discrepancies for all groups.
5. To the extent that housing market discrimination accounts for SP-RH mismatches, controls for inertial and spatial aspects of migration, as well as socioeconomic sorting, will fail to account for SP-RH differences. Furthermore, these persistent discrepancies should be greatest for the most negatively stereotyped groups, namely Blacks and Latinos.

2.3 Data: Stated Preferences and Residential History in Los Angeles County

I use data from the Los Angeles Family and Neighborhood Survey (L.A.FANS) to examine the above hypotheses. Uniquely, this survey contains both SP and RH reports. L.A.FANS began as a stratified, clustered sample of Los Angeles County households interviewed in the years 2000 to 2002.² Within households, randomly selected adult respondents (RSAs) reported on two years of residential history preceding the Wave 1 interview, among a number of topics. Beginning in 2006, RSAs were re-interviewed. In this second wave of data collection, RSAs who remained in Los Angeles County reported on their residential histories in the intervening years between the Wave 1 and the Wave 2 surveys. In addition to a re-interview of the original respondents, the second wave of L.A.FANS added a replenishment sample of

²See Peterson et al. (2004, 2012) and Sastry et al. (2006) for more details on the sampling procedures.

new RSAs. These RSAs completed residential preferences instruments and reported their residential histories for the six years preceding the interview.³

Key to testing the hypotheses I outline above, I classify L.A.FANS respondents by their racial identifications. I classify RSAs according to a four category scheme. I distinguish between Latinos, non-Hispanic Whites, non-Hispanic Blacks, and non-Hispanic Asian and Pacific Islanders. I combine the seven respondents who reported other races and ethnicities into the White category. I assign respondents to these four categories based on respondents' own reports or the reports of another household member for the few respondents who did not identify.

In addition to the residential history reports, Wave 2 of L.A.FANS prompted RSAs to state their neighborhood preferences using two SP instruments. The first preference instrument asked respondents to rank a set of hypothetical neighborhoods which differed in their racial compositions. The second preference instrument asked respondents about the racial composition of their ideal neighborhoods. My analytic sample comprises the 1,214 RSAs who provided RH reports and responses to both SP instruments. Below, I discuss the preference instruments and basic descriptive results based on these instruments. I then discuss the RH data. In discussing the RH data, I also discuss the United States Census data I use to generate neighborhood level variables for the RH analysis, as well as the individual-level variables, drawn from L.A.FANS, that could influence RH neighborhood outcomes.

2.3.1 Stated Preference Data in L.A.FANS

The L.A.FANS survey assessed new and panel respondents' stated racial preferences using two variations of a showcard vignette first implemented in the Detroit Area Study (Farley et al. 1979, 1978) and then deployed across a number of metropolitan contexts in the Multi-City Study of Urban Inequality (Farley et al. 1997). In the first variation of the experiment, respondents were presented with five images, each depicting a hypothetical neighborhood

³Wave 1 RSAs who left Los Angeles County were also interviewed at Wave 2, but they completed a shorter survey that did not assess residential history or preferences. Thus my sample only includes those persisting in Los Angeles County as of the Wave 2 interview.

as in Figure 2.1. Each house in each hypothetical neighborhood was filled in to indicate whether an Asian, Black, Latino, or White family resided there. Interviewers then prompted each respondent to rank these neighborhoods in order of where they would most like to live.⁴ A random algorithm determined the racial compositions presented to respondents, including a single neighborhood that reflected the racial composition of respondents’ neighborhoods at the time of the survey.⁵

In the second version of the experiment, interviewers prompted respondents to fill out a blank card, like that in Figure 2.2, to indicate the racial mix in the “ideal” neighborhood where they would want to live.⁶ In this ideal vignette, respondents could choose any of 680 possible neighborhood racial compositions, even racial compositions that corresponded to no extant neighborhoods in Los Angeles County.⁷ For both the ideal neighborhood response and the ranked neighborhood vignettes, I characterize each hypothetical neighborhood according to four racial composition components: proportion Asian, proportion Black, proportion Latino, and proportion White.

Table 2.3 contains summary statistics for L.A.FANS respondents SP choices. Panel A of Table 2.3 describes the average racial compositions of respondents’ ideal neighborhoods, averaged by respondent race. For ideal neighborhood composition, only Whites, on average, express a mean preference for neighborhoods with a majority of own group residents. Besides

⁴Verbatim: “Imagine that you were looking for a place to live. You found nice, affordable places in five different neighborhoods. The neighborhoods have different numbers of White, Black, Asian, and Latino families. I’ll show you drawings of these neighborhoods on the computer screen in a minute. Please tell me which one would be your first choice as a place to live, your second choice, and so on.”

⁵The racial compositions presented to respondents, while varying across cards, were skewed towards neighborhoods similar to those in which they already resided. That is, the set of vignettes do not appear to constitute a fully factorial experiment of the form advocated by (Louviere et al. 2000). This has advantages and disadvantages. On one hand, respondents were more likely to be prompted with plausible neighborhood alternatives, and so may have had a better time of imagining these neighborhoods. On the other hand, because of similarity between neighborhoods in the ranking experiment, respondents may have found it difficult to distinguish between neighborhood alternatives, potentially introducing more error into their responses.

⁶Verbatim: “Now imagine your ideal neighborhood that has the ethnic and racial mix that you would personally feel most comfortable living in. Here is a blank card like those I showed you on the screen. Please put a letter in each house on the card to show your ideal neighborhood where you would most like to live.”

⁷Given 14 empty houses, as in Figure 2.2, and four groups, there are $\binom{14+4-1}{4-1} = 680$ possible racial compositions, ignoring the spatial arrangement of groups into houses.

preferring an own-group majority, Whites' ideal neighborhoods have roughly equal shares of Asians, Latinos, and Blacks, albeit with a slight preference for Latino and Asian neighbors over Black neighbors. Both Asian and Latino respondents express an average preference for own-group plurality neighborhoods in the ideal SP vignette, with Latinos exhibiting a stronger in-group preference. Respondents identifying as Asian and Latino also express a preference for some mixing with Whites—Whites are the second largest group in their ideal neighborhoods. Overall, non-Blacks are reluctant to include Blacks in their ideal neighborhoods, with Blacks making up the smallest proportions of Latinos', Asians', and Whites' ideal neighborhoods. Meanwhile, Blacks' ideal neighborhoods have a joint Black-White majority, with roughly equal mixing between Black and White residents (31%), with lower, but balanced, representation of Asian and Latino residents.

Panel B of Table 2.3 presents summary statistics for respondents' first ranked neighborhoods and the full set of neighborhoods presented to respondents in the rank SP instrument. The results from the rank SP instrument cannot be interpreted directly, because respondents' choices were dictated by the set of neighborhoods presented to them. And respondents from different racial groups were not presented with the same hypothetical neighborhoods in the rank SP vignette. Instead, White respondents were presented with neighborhoods that were, on average, predominantly White, Latinos were presented with predominantly Latino neighborhoods, and Blacks and Asians with neighborhoods that were disproportionately Black and Asian, respectively. These data can only be productively interpreted using methods that account for the availability of different neighborhood types in the rank SP vignettes. I discuss these methods later on.

2.3.2 RH: Combining Individual-Level Data in L.A.FANS with Neighborhood-Level Data from the Census

The RH data in L.A.FANS catalogue RSAs residential locations over the period 1998-

2008.⁸ I discretize the residential history into person-quarters and examine neighborhood locations at the end of each person-quarter. I discretize time by quarters, rather than years, in order to capture as many moves as possible among respondents who reported more than one move in some years. RSAs' residential histories are geocoded according to 2000 Census tract boundaries, which I use to link individuals to data about their neighborhoods. I include a person-quarter in the analysis if an individual started and ended the interval in a valid Los Angeles Census tract.⁹

To characterize the racial composition of Los Angeles County neighborhoods, and thus the neighborhoods of L.A.FANS respondents, I combine data from the 1990, 2000, and 2010 United States Census. Data for 1990 and 2000 are aligned to 2000 Census tract boundaries in the Neighborhood Change Database (NCDB) produced by the Urban Institute and Geolytics (Geolytics 2003). I align 2010 data to 2000 tract boundaries using tract relationship files provided by the Census and use linear interpolation to estimate the composition of tracts in intervening years, 1998-2008, for which adult respondents provided residential history. I distinguish between four groups, non-Hispanic Asian or Pacific Islander, non-Hispanic Black, Latino, and non-Hispanic White, in characterizing each neighborhood's racial composition. I group the few Los Angeles residents who identified with other groups on their Census forms with Whites.

I use data from the 1990 and 2000 Census long forms, normalized to 2000 tract boundaries in the NCDB, in combination with data from the 2005-2009 and 2006-2010 American Community Surveys to characterize the socioeconomic characteristics of neighborhoods. I consider three main socioeconomic attributes of neighborhoods: income, education, and home ownership. I expect L.A.FANS respondents to move into neighborhoods where their income matches the income of other residents, primarily because of housing costs that are highly

⁸Over 60% of the Wave 2 interviews, and thus nearly all the residential history observations, precede the December 2007 onset of the Great Recession. Because most observations precede the recession, I do not give careful consideration to how the collapse of housing prices influenced subsequent residential decisions.

⁹I set this condition because L.A.FANS did not provide geocodes in most cases when individuals reported an address outside of Los Angeles County. It was unclear how to construct the neighborhood choice sets for respondents moving to Los Angeles from outside the County (see methods discussion below).

correlated with neighborhood income. To characterize the income “fit” between L.A.FANS respondents and Los Angeles neighborhoods, I estimate the household income distributions within each neighborhood. I obtain estimates of the income distribution by fitting gamma distributions to binned income data.¹⁰ I combine these estimated income distributions with data on L.A.FANS respondents’ incomes to derive a percentile figure, ranging from zero to one, indicating where each L.A.FANS respondent in my sample fits into each prospective destination neighborhood’s income distribution. A value of zero indicates that the respondent falls at the very bottom of the neighborhood income distribution, and a value of one indicates that the respondent sits at the top of the income distribution. A value of .5 indicates the respondent’s family’s income matches the (estimated) median household income for the neighborhood.

Beyond income, levels of education likely influence residential choices, perhaps especially among the college educated. Those with college educations might seek out co-residence with other college graduates because of shared educational values or social networks that either develop during college, or during careers that people select into based on college attendance. Thus, I characterize each neighborhood by the percentage of those over age 25 who have received a bachelors degree or more. Finally, home ownership should be highly influential in inter-neighborhood migration, not least because those who want to own a home can only do so in a neighborhood that has homes available for purchase. In addition to this mechanical effect, potential home owners may have economic or ideological preferences to live in neighborhoods where other residents are home owners as well. Summary statistics for Los Angeles County Neighborhoods are provided in Table 2.1.

Table 2.4 summarizes RH outcomes for L.A.FANS respondents. Latinos live in predominantly Latino neighborhoods, with Whites representing the second largest group, followed

¹⁰I attempted to fit several different distributions, including a log-normal and Wiebull distribution, to the neighborhood income data. Among the distributions I tested, the gamma distribution provided the best overall fit across all neighborhoods, in terms of fit statistics like BIC, as well as relative to the reported median and means of the income distribution. The mean and median values obtained from these estimates all came very close to matching the observed mean and median incomes.

by Asians and Blacks. Latinos live in neighborhoods with the lowest household incomes, the lowest representations of bachelor's degree holders, and low levels of home ownership. Even in these low socioeconomic status neighborhoods, the family incomes of Latino respondents fall below the neighborhood median household income. Whites live in neighborhoods that are majority White, with substantial Latino representation and very few Blacks. Whites' neighborhoods have the highest median household incomes, rates of home ownership, and prevalence of bachelor's degrees. Blacks, on average, live in neighborhoods that are plurality Latino, with Whites the second largest group, out-numbering even Blacks. Asian L.A.FANS respondents live in neighborhoods that are roughly 35% Asian and 33% White, with Latinos making up the third largest group. These results are qualitatively similar to results obtained using U.S. Census Data for L.A. County, shown in Table 2.1. However, the Black respondents in L.A.FANS appear exceptional, as their exposures to Black neighbors are lower, and their exposures to White neighbors higher, on average, than observed for Blacks in Los Angeles County as a whole.

Table 2.4, Panel C also summarizes rates of residential mobility in the L.A.FANS sample. Overall, the 1,214 L.A.FANS respondents contribute 898 person-quarter mobility events, with nearly half the sample (600 respondents) reporting at least one residential move within Los Angeles County over the survey period. Weighted, this corresponds to 0.074 moves per year.¹¹ Blacks and Latinos have relatively high rates of mobility, while Asians and Whites have lower rates of residential mobility.

¹¹This is less mobility than reported for persistent Los Angeles County residents in the 2005-2010 American Community Survey, in which approximately 10% of L.A. County persisters, ages 18 or older, reported making moves between houses within L.A. County in the prior year. However, the L.A.FANS respondents are relatively older than the Los Angeles County population, with the preponderance having been drawn into the sample conditional on being age 18 or older in 2000-2002. The lack of younger respondents in the replenishment sample may have rendered it difficult to weight the sample appropriately by age. Notably, the observed mobility rates are consistent with those observed for year-to-year Los Angeles County persisters ages 30 and older.

2.3.3 RH: Individual Level Controls

Migration decisions are not made solely based on neighborhood-level factors. People also seek out housing and neighborhood options that either match their needs, or fall within the boundaries imposed by individual and family level constraints. In the models presented here, these constraints are represented as interactions between peoples' resources and prior experiences and the characteristics of neighborhoods. So, in addition to racial identification, I consider a number of non-racial individual characteristics that may influence RH outcomes. I focus on three individual-level socioeconomic factors that I pair up with the previously discussed neighborhood-level factors: income, education, and home ownership.

I account for the likelihood that people will move into neighborhoods that their income affords using measures of family income obtained in Wave 1 and Wave 2 of L.A.FANS. I rely on the family income reports rather than employment histories partly out of an expectation that family income, rather than individual or household income, shapes prospective housing choices. But I also face a data limitation—L.A.FANS did not obtain complete household income reports, instead opting to ascertain family incomes for RSAs. This income includes an RSA's spouse, cohabiting partner, and children, where applicable. Roommates and other household residents are omitted from the income figures. By ignoring the potential contributions of roommates and other household residents, I assume that their incomes did not factor into RSAs' residential choices. Full family income reports for L.A.FANS respondents were obtained only for the years corresponding to the Wave 1 and Wave 2 interviews. Only partial reports, based on respondents' employment histories, were obtained for intervening years. L.A.FANS obtained income reports in components, distinguishing between earnings, investment income, welfare receipt, retirement income, and so on. Where respondents' component reports were missing, L.A.FANS used an imputation procedure to estimate the component of income, relying only on data from elsewhere in the income survey as well as the household

roster.¹² In Wave 1, slightly less than 30% of respondents in my sample had some component of income imputed. In Wave 2, slightly less than 40% of respondents in my sample had at least one component of income imputed. These figures overstate the degree of missingness, however, because they include missingness across all components and family members.

For those who had family income reports in both Wave 1 and Wave 2 surveys, I average the Wave 1 and Wave 2 income reports to estimate respondents' yearly financial resources. By construction, new entrant RSA respondents in Wave 2 did not have Wave 1 income reports. To estimate their average income over the preceding six year period corresponding to their residential history reports, I estimate a Wave 1 income by depreciating their Wave 2 incomes using a regression of Wave 1 income on Wave 2 income, education, race, gender, age, marital status, and survey year among those with both Wave 1 and Wave 2 income reports. I then averaged this depreciated income with the observed Wave 2 income. I adjust all income figures to 1999 dollars.

To account for the sorting of individuals into neighborhoods based on educational attainment, I distinguish between L.A.FANS respondents who completed at least a bachelor's degree, and those who did not. To account for the matching of home owners to neighborhoods where home ownership is common, I distinguish between respondents who owned their homes at the Wave 2 survey, and respondents who reported renting at Wave 2.

The summary statistics for L.A.FANS respondents' individual and family characteristics are presented, by race, in Table 2.2. In accordance with Census data and previous research on socioeconomic status of Los Angeles residents, Latinos and Blacks are the most socioeconomically disadvantaged groups, with the lowest family incomes, the lowest levels of college completion, and the lowest levels of home ownership. Black and Latino respondents were

¹²Missingness for an income component could occur for two reasons. Either the respondent did not say whether a given family member had received income from a given source, or the respondent claimed income from the source, but could not recall the amount.

also younger and less likely to be married than Whites and Asians.¹³

To make inferences about the relative effects of racial composition and non-racial factors on RH outcomes, and to properly compare SP and RH data in Los Angeles, I combine these individual-level data and neighborhood level data using a discrete-choice modeling framework. This framework explicitly accounts for the opportunity structure individuals face when moving between neighborhoods, and allows me to account for the matching of individuals to neighborhoods along both racial and non-racial lines. I discuss these methods briefly in the following section. A more in depth discussion is provided in Chapter 3.

2.4 A Discrete Choice Approach to combining SP and RH Neighborhood Choice Data

I use discrete choice logistic regression models to investigate SP and RH differences in Los Angeles. In short, the models are multinomial logistic regression models that allow for the inclusion of alternative specific regressors. Respondents, indexed by i , make neighborhood choices across multiple situations, indexed by t . In each choice situation, respondents are faced with a *choice set* of discrete, mutually exclusive neighborhoods, denoted by C_{it} . The choice set could differ across respondents and across choice situations. Discrete choice models are random utility models that assume that each person, in each choice situation, assigns a utility, U_{ijt} , to each neighborhood j in his choice set.

$$U_{ijt} = V_{ijt} + \epsilon_{ijt} \tag{2.1}$$

The “observed” part of the utility, V_{ijt} , is parameterized according to the observed characteristics of neighborhoods, potentially in interaction with the respondents’ own characteristics.

¹³I weight these and other individual-level summary statistics using cross-sectional weights provided for panel and new entrant RSAs in Wave 2 of L.A.FANS. These weights are adjusted to account for attrition between Wave 1 and Wave 2 surveys, and are intended to render the sample representative of Los Angeles County adults at the time of the Wave 2 survey.

The unobserved aspects of choices are captured in the random disturbance, ϵ_{ijt} .

The models assume that respondents choose the neighborhood that provides the highest total utility. By assuming the unobserved parts of the utility, ϵ_{ijt} , are uncorrelated with the observed variables and that they are drawn independently from a standard Gumbel distribution, the probability that individual i chooses neighborhood j in situation t is:

$$P_{ijt} = \frac{\exp(V_{ijt})}{\sum_{k \in C_{it}} \exp(V_{ikt})} \quad (2.2)$$

Multiplying these probabilities across a sample providing data on SP and RH choices yields a likelihood function whose parameters can be estimated using traditional maximum likelihood techniques. Chapter 3 contains details about how to construct the likelihood function.

I parameterize the “observed” part of the utility according to racial composition components, and additional non-racial components for the RH data, as follows:

$$V_{ijt} = \beta_{is}X_{jt} + \zeta_{it}Z_{ijt} \quad (2.3)$$

X_{jt} is a column vector of observed racial attributes, like proportion Black or proportion Latino. β_{it} is a row vector of coefficients which denote the weight given to the racial attributes of neighborhoods in the decision process. The i subscript for β indicates that the effects could differ across individuals, and the t subscript indicates the effects could differ across different choice situations. Z_{ijt} is a column vector of non-racial neighborhood characteristics. These could be objectively evaluated, like neighborhood poverty, median income, age composition, etc. Or they could be subjectively evaluated, like the distance of a neighborhood from an individual’s place of work, or whether or not a close family member is residing in the neighborhood. ζ_{it} is a row vector of coefficients which indicates the influence of non-racial aspects of neighborhoods in residential choices. The subscripts indicate that these coefficients could differ across situations and across agents.

In the L.A.FANS, respondents provide data on neighborhood choices across three different

kinds of choice situations: Ideal SP, rank SP, and RH. I alternately refer to these different sets of situations as *scenarios*. As previously mentioned, Chapter 3 provides additional details about how to combine data from multiple sources. The ideal SP data provide information on a single choice situation, with respondents identifying one ideal neighborhood racial mix among 680 possible racial mixes. I parameterize the ideal SP utility functions for each neighborhood, k , in the 680 neighborhood choice set as follows:

$$\begin{aligned}
 V_{ik}^{SPI} = & \beta_{i1}^{SPI} LATINO_k + \beta_{i2}^{SPI} LATINO_k^2 + \beta_{i3}^{SPI} BLACK_k + \beta_{i4}^{SPI} BLACK_k^2 \\
 & + \beta_{i5}^{SPI} ASIAN_k + \beta_{i6}^{SPI} ASIAN_k^2 + \beta_{i7}^{SPI} WHITE_k^2
 \end{aligned} \tag{2.4}$$

LATINO, *BLACK*, *ASIAN*, and *WHITE* terms denote the representation of each group, recorded as proportions, in the hypothetical neighborhood. I include quadratic terms to account for potential resistance to neighborhoods with very high or very low representation of some groups.¹⁴ I index the β 's by i to indicate that the effects of racial composition will differ across respondents. In particular, I assume that racial composition preferences differ across four racial groups. I accomplish this by estimating separate models for each group. The *SPI* superscript distinguishes these racial composition effects from the effects in the rank SP situations, discussed below. In other words, I allow for the possibility that racial composition preferences differ across ideal SP, rank SP, and RH situations.

In the rank SP situations, respondents rated five neighborhoods. I “explode” these rank SP data into four choice situations—one situation from the ranking of the first neighborhood above all others, another from the ranking of the second ranked neighborhood above all neighborhoods except for the first, and so on.¹⁵ I parameterize the utility for the r^{th} ranked

¹⁴I also tested models that included third order polynomials in racial composition. These models tended to improve model fit, but inspection of the results suggested that these third order effects were fitting heterogeneity within racial groups that would be better dealt with by using methods that allow for unobserved heterogeneity. Future versions of this work will employ these models to explore heterogeneity in racial composition preferences.

¹⁵Around 10% of respondents had at least two neighborhoods that were tied in their rankings. In separate models, I used an exact technique to account for these ties (Allison and Christakis 1994). The results from these “exact” models generally did not differ from the results I obtained by randomly breaking ties within respondents. The results I present here are based on randomly breaking ties.

neighborhood as I did with the ideal SP situations.

$$\begin{aligned}
V_{ir}^{SPR} = & \beta_{i1}^{SPR} LATINO_r + \beta_{i2}^{SPR} LATINO_r^2 + \beta_{i3}^{SPR} BLACK_r + \beta_{i4}^{SPR} BLACK_r^2 \\
& + \beta_{i5}^{SPR} ASIAN_r + \beta_{i6}^{SPR} ASIAN_r^2 + \beta_{i7}^{SPR} WHITE_r^2
\end{aligned} \tag{2.5}$$

I assume that the racial composition effects do not differ across rankings.¹⁶ The *SPR* superscript indicates that these effects may be different from the ideal SP effects.

Finally, each respondent faces multiple RH choice situations, corresponding to each person-quarter of residential history. In RH scenarios, I assume that respondents choose one neighborhood from all the neighborhoods in Los Angeles County. I parameterize the utility that respondents assign to neighborhoods as follows:

$$\begin{aligned}
V_{ijt}^{RH} = & \beta_{i1}^{RH} LATINO_{jt} + \beta_{i2}^{RH} LATINO_{jt}^2 + \beta_{i3}^{RH} BLACK_{jt} + \beta_{i4}^{RH} BLACK_{jt}^2 \\
& + \beta_{i5}^{RH} ASIAN_{jt} + \beta_{i6}^{RH} ASIAN_{jt}^2 + \beta_{i7}^{RH} WHITE_{jt}^2 \\
& + \zeta_{i1} LNUNITS_{jt} + \zeta_{i2} PROXIM_{ijt} + \zeta_{i4} \sqrt{DIST_{ijt}} \\
& + \zeta_{i4} INCPOS_{ijt} + \zeta_{i4} INCPOS_{ijt}^2 \\
& + \zeta_{i5} PCOLL_{jt} + \zeta_{i6} POWN_{jt} \\
& + \zeta_{i7} ORIG_{ijt} \times AGE_{it} + \zeta_{i8} ORIG_{ijt} \times AGE_{it}^2 \\
& + \zeta_{i9} ORIG_{ijt} \times OWNER_i + \zeta_{i10} ORIG_{ijt} \times MARRIED_i \\
& + \zeta_{i11} SPLANG_{jt} + \zeta_{i12} ASLANG_{jt}
\end{aligned} \tag{2.6}$$

The racial composition terms are as in the SP models, but with that data for neighborhood racial composition taken from the Census, rather than given in the L.A.FANS data. The *RH* superscript signals that these coefficients may be distinct from the ideal and rank SP coefficients, while the *i* subscript indicates that these coefficients may differ across respondents, i.e. according to racial identification. In addition to these main effects, I also include

¹⁶In separate models using only the rank SP data, I tested for consistency of effects across ranks. I generally could not reject the null hypothesis of no difference in preferences across rankings.

interactions between the racial composition terms and a dummy variable that identifies the alternative in the choice set that corresponds to the respondent’s own house. This accounts for the possibility that racial composition may influence choices to move differently than choices of where to move.

Unlike the ideal SP and rank SP choice situations, observable non-racial features of neighborhoods influence RH choices. The key analytic approach taken here, as with much research on racial and ethnic differences in neighborhood outcomes, is to add in non-racial variables in the analyses as potential mediators of RH and SP differences. Do statistically significant differences between β^{RH} , β^{SPI} , and β^{SPR} persist after introducing controls for non-racial characteristics of neighborhoods?

I include a number of non-racial controls as potential mediators. $LNUNITS$ is the log of number of housing units in neighborhood j in quarter t . All else being equal, and holding the vacancy rate constant, neighborhoods with more housing units should have more vacancies in each period, and will be more likely to receive movers.

$PROXIM$ and $\sqrt{DIST}$ address the unobserved costs related to the decision to migrate and the distance migrated. The low mobility rates in the quarterly L.A.FANS residential histories reveal that respondents are very likely to “choose” their own houses from quarter to quarter. I also expect respondents to be less likely to move to tracts that are more geographically distant. I parameterize this with a series of dummy variables distinguishing between alternatives in the Los Angeles County choice set that are the respondent’s own house, a new house in the respondent’s origin neighborhood, a new house in a first, second, third, or fourth order adjacent neighborhood, or a new house in fifth or higher order adjacent neighborhood. $\sqrt{DIST}$ is the square root of the distance, measured in miles, between the centroid of a respondent’s origin tract and the the centroid of a potential destination tract.¹⁷

The $INCPOS$ terms identify the position of the respondent’s family in the neighborhood income distribution. I include linear and quadratic effects because people will likely avoid

¹⁷This distance is “as the crow flies.” Future versions of this study will use transit times to better represent the spatial accessibility of neighborhoods.

neighborhoods where they are either much richer or much poorer than other neighborhood residents.¹⁸ Because of the limited nature of the L.A.FANS income reports, I treat family income as a time invariant variable.

PCOLL and *POWN* control for tendencies, among some respondents, to move to neighborhoods based on the availability of homes to own, rather than rent, and the presence of highly educated residents. These variables respectively indicate the proportion of neighborhood j 's residents who have bachelor's degrees, and the proportion of occupied housing units that are occupied by owners. I assume that the effects of these variables differ across respondents based on their own educational attainments and housing tenures. I interact *PCOLL* with an indicator of whether a respondent attained a bachelor's degree by Wave 2. I expect those with college degrees to be more likely to move to neighborhoods where others with college degrees live. This could be due to preferences or due to social networks that are segregated by educational attainment. I interact *POWN* with an indicator of whether the respondent ever reported owning a home, as those who own homes, whether due to preferences or necessity, will likely live in neighborhoods with higher levels of home ownership. Additionally, I account for selection into the inter-neighborhood migration stream by including interactions between the origin house identifier, *ORIG*, and respondent-level age (*AGE*), home ownership status (*OWN*), and marital status (*MARRIED*) variables. I expect greater immobility among those who are older, own a home, or are married.

Finally, there may be a disconnect between the racial ascription respondents engage in when evaluating neighborhoods, and the processes of racial identification that produced the Census data I use to describe Los Angeles neighborhoods. It is unlikely that respondents in L.A.FANS ascriptively identify each self-identified Latino neighbor as Latino, or each Asian neighbor as Asian, or each Black neighbor as Black. Housing seekers may use behavioral,

¹⁸I examined several measures and specifications to account for sorting on income, including specifications based on the ratio between respondents' incomes and the neighborhood median income, the difference between these incomes, and multiplicative interactions between these incomes. I tested logs, splines, and polynomials in these measures. In the end, the second order polynomial specified in the main text provided the best and most parsimonious fit.

visual, and linguistic cues to do this ascriptive work. I expect these cues to be especially important for ascribing racial categories to self-identified Latino and Asian neighbors, groups which fall in the middle of the Los Angeles racial hierarchy and whose members vary in their nativity status, national origins, and socioeconomic attainments (Charles 2006; Treiman and Lee 1996). I proxy these cues with the linguistic variables *SPLANG* and *ASLANG*. I generate these variables by dividing the number of neighborhood residents who speak Spanish or an Asian language at home, and who speak English less than very well, by the number identifying as Latino or the proportion identifying as Asian, respectively. Generally, these variable fall between 0 and 1, with 0 signaling a high degree of English language proficiency among self-identified Asians and Latinos, and 1 indicating a low degree of English language proficiency. In alignment with the racial composition effects, I include second order polynomials for these variables.¹⁹

As discussed in Chapter 2, I pool the RH, rank SP, and ideal SP likelihood components into a single likelihood function. I estimate the pooled model using Stata’s “clogit” command (StataCorp 2011). I use Stata’s cluster feature to adjust the variance-covariance matrix based on the multiple SP and RH observations contributed to the analysis by each respondent. I estimate sets of models using a hierarchical, incremental approach, adding in sets of non-racial controls to examine whether and to what degree these variables act as mediators that “explain” the difference between racial composition coefficients across SP and RH choice situations.

¹⁹I include one other control in the Ideal SP and RH portions of the model that have some technical importance. The possibility of choosing among 680 neighborhoods in the Ideal SP situation, and over 2,000 Census tract in the RH situations, this across 1,214 respondents who contribute approximately 28 person-quarter RH observations each, poses a challenge in computation. Fortunately, conditional logistic regression models permit sub-sampling of the choice sets within respondents. I take a stratified sample of tracts within respondents in each RH choice scenario. I include the chosen tract, the origin house, and the origin neighborhood in the analytic choice set with 100% probability. I randomly include an additional 97 alternatives (100 alternatives in total) in each respondent’s choice set, giving preference to first through fourth order adjacent tracts. In the ideal SP data, I sample the chosen tract with 100% certainty, and then randomly include an additional 49 alternatives (50 alternatives total) in each respondent’s ideal SP choice set. Parameter estimates are consistent if the log of the sampling fraction, within choice situations, is included as a regressor with its coefficient constrained to negative one (Ben-Akiva and Lerman 1985; Bruch and Mare 2012).

2.5 Model Results

Racial composition coefficients from the full model, including all non-racial controls, are contained in Table 2.5. Table 2.6 contains the estimated coefficients for non-racial variables. These coefficients are difficult to interpret by themselves. Instead, I interpret the model output based on three post-estimation steps. First, I produce sets of predicted probabilities of neighborhood choice as functions of racial composition, separately for Latino, White and Black Respondents.²⁰ I then plot these predicted probabilities for a qualitative view of the degree and direction of SP-RH mismatch for members of these groups. I pair these plots with Wald tests of racial composition coefficients from the models, paying special attention to statistical tests of differences between SP and RH situations. Second, I use SP and RH coefficients to calculate predicted neighborhood choice probabilities for “typical” Latino, White, and Black respondents choosing among neighborhoods whose racial compositions match those of neighborhoods in Los Angeles. I calculate indices of dissimilarity to summarize the differences between SP and RH choice probabilities. Third, I calculate expected neighborhood racial compositions based on these L.A. choice set predicted probabilities.

2.5.1 Predicted Probabilities and Wald Tests, by Respondent Race

Figures 2.4 through 2.14 depict predicted probabilities of neighborhood choice for Latino, White and Black residents of Los Angeles. I generate predicted probabilities based on a simulated choice set of 100 house neighborhoods. This choice set contains one neighborhood for each unique racial composition achieved by apportioning 100 houses among families falling into four racial categories. Using the relevant racial composition coefficients, and holding all other covariates constant, I calculate a choice probability for each neighborhood in this hypothetical choice set. To plot probabilities as function of the four racial composition components—proportion Latino, Black, White, and Asian—I calculate the mean choice

²⁰I omit discussion of Asian respondents, as the sample of Asians in L.A.FANS was small and idiosyncratic when compared to Los Angeles County as a whole.

probability at each value of one of the racial composition components.²¹

In addition to these predicted probability plots, I also test the statistical significance of racial composition effects. I examine whether and which racial composition components have statistically significant effects *within* ideal SP, rank SP, and RH situations. Key to my hypotheses, I also test for significant differences *between* ideal SP and RH, and between rank SP and RH coefficients. I use cluster adjusted Wald tests to perform these tests. Tests of within situation significance are shown in Table 2.7. Tests of between scenario differences are displayed in Table 2.8. I first present results for Latinos, followed by Whites and Blacks.

2.5.2 Latino Respondents

Latinos in Los Angeles do pay attention to racial composition when making residential choices. In RH scenarios, proportion Asian, Black, and White have statistically significant effects on choices (using $p = 0.05$ as a threshold), while the effect of proportion Latino is marginally significant (Panel A, Column 1 of Table 2.7). Likewise, Latinos express significant racial preferences as well. In their SP responses, Latinos expressed preferences across all racial composition components considered in L.A.FANS: proportion Asian, Black, Latino, and White (Panel A, Columns 2 and 3 of Table 2.7).

Latinos also endure statistically significant disconnects between SP and RH, even when controlling for a full set of non-racial factors in RH models. Table 2.8 Panel A shows that when I consider all racial composition terms simultaneously, I find statistically significant differences between ideal SP and RH choices (Column 1) and rank SP and RH choices (Column 2). However, the results based on ideal and ranked stated preferences are not in statistical agreement. Only the effect of proportion White is statistically different between rank SP and RH scenarios (Column 2), but this is the only effect that is not significantly different between ideal SP and RH scenarios (Column 1).

²¹An average is necessary because there are many neighborhood configurations that could be consistent with a given value of a racial composition component. For example, when generating a mean choice probability for neighborhoods with 20% Black composition, I calculate a mean over all neighborhoods that are 20% Black, but which vary in Asian, Latino, and White representation in the 0-80% range.

I predicted that Latinos, as a negatively stereotyped and socioeconomically disadvantaged group in Los Angeles, would have greater exposure to Latinos and Blacks than they prefer, and less exposure to Whites. To evaluate the qualitative aspect of these hypotheses, I examine SP and RH choice probabilities as functions of proportion Latino, White, Black, and Asian.

Figure 2.3 shows that Latinos have a preference for more Latino neighbors. Across all SP and RH scenarios, and regardless of the controls employed in the RH portion of the model, Latinos are more likely to choose a neighborhood as the proportion Latino increases. RH and rank SP choice probabilities largely agree, once I account for the proximity of destination neighborhoods. The main difference crops up between these results and the ideal SP results. The ideal SP data suggest a significantly stronger preferences for Latino neighbors (Table 2.8, Panel A, Column 1). The difference between rank SP and RH coefficients, in contrast, is not significant at any reasonable threshold. Rather than being over-exposed to Latino neighbors, as I predicted, Latinos are either exposed in accordance with their preferences, or move to neighborhoods with fewer Latino neighbors than they claim to prefer.

The story for proportion Black is similar. Latinos state greater unwillingness to choose a neighborhood as the proportion Black increases, as shown in Figure 2.4. 0% Black neighborhoods have the highest choice probabilities, and 100% Black neighborhoods have the lowest choice probabilities in both SP scenarios. Latinos' RH neighborhood choices also signal resistance to Black neighbors, although with indifference over the 0-30% Black range. According to Table 2.8, Panel A, the difference in the effects of proportion Black on neighborhood choice is not significantly different between RH and rank SP scenarios. Again, the statistically significant difference is between ideal SP and RH outcomes. In the ideal SP data, Latinos express a stronger resistance to Black neighbors than is evinced in rank SP and RH situations. Only if we give credence to the ideal SP data over the rank SP data would we conclude that Latinos tend to move to neighborhoods with more Blacks than they prefer. Overall, however, the differences are mainly of degree, and not of kind.

The results for proportion Asian resemble the results for proportion Black. The relevant predicted probabilities are plotted in Figure 2.6. RH and rank SP predicted probabilities basically agree, as confirmed by the Wald test in Panel A of Table 2.8. Latinos prefer and move to neighborhoods that are between 0 and 40% Asian, and avoid neighborhoods that are more than 40% Asian. Once again, the ideal SP results differ more in degree than in kind. In the ideal SP data, Latinos express sharper resistance to neighborhoods as the proportion Asian increases. I can only accept that Latinos are unable to match RH and SP with respect to proportion Asian if I assume that the ideal SP data represent Latinos “true” preferences better than the rank SP data.

The results for proportion White diverge from the preceding patterns. Figure 2.5 shows that Latinos prefer neighborhoods with intermediate to high levels of Whites in the rank SP vignette, and are more likely to pick neighborhoods with 100% White residents than they are to pick neighborhoods with 0% White neighborhoods. However, in the both the ideal SP and RH data, Latinos were more likely to choose or move to neighborhoods with low to intermediate representation of Whites, with higher choice probabilities for 0% White neighborhoods than for 100% White neighborhoods. Table 2.8, Panel A, shows that the rank SP-RH difference in the effect of proportion White is statistically significant, while the ideal SP-RH difference is not statistically significant. The rank SP data would lead me to confirm my hypothesis that Latinos are underexposed to Whites relative to their preferences, even when accounting for non-racial factors that should influence actual migration, whereas the ideal SP data would lead me to reject this hypothesis. As I will discuss later on, this ambiguous result falls into starker relief when I calculate expected neighborhood racial compositions based on a choice set of Los Angeles neighborhoods.

2.5.3 White Respondents

Both SP and RH data sources show that Los Angeles County Whites respond to racial composition when making neighborhood choices, as seen in Table 2.7 Panel B. In RH data

(Column 1), Whites showed sensitivity to the presence of Asians, Blacks, and Latinos ($p < 0.05$), although only marginally significant sensitivity to the presence of Whites over and above that implied by the other racial composition components ($p < 0.10$). Results from the SP scenarios (Columns 2 and 3) show sensitivity across all racial composition components ($p < 0.05$), with that sensitivity only marginally significant for proportion Asian and Latino in the ideal SP data ($p < 0.10$).

Whites are the most advantaged group in LA County. They have the highest socioeconomic attainments, and face few negative stereotypes (Charles 2006; Bobo and Johnson 2000). I hypothesized that, as an advantaged group, Whites would be more likely than minority group members to obtain a match between their stated preferences and their residential experiences. Table 2.8, Panel B presents results from Wald tests of these hypotheses. Overall, I can reject the null hypothesis that Whites' racial composition preferences match their actual neighborhood choices. However, the results differ across groups and scenarios. I observe statistically significant differences between ideal SP and RH effects for each separate racial composition component, but only one significant difference, for proportion Asian, between rank SP and RH scenarios. Beyond achieving an SP-RH match, I also expected Whites to both prefer and reside in neighborhoods with high White representation, and to live in neighborhoods with few disadvantaged minorities, particularly Blacks and Latinos. For a qualitative assessment of these hypotheses, I turn to the predicted probabilities in Figures 2.7 through 2.10.

The predicted probabilities in Figure 2.7 suggest that Whites are generally resistant to Latinos as neighbors. However, Whites' actual neighborhood choices differ slightly, in substantive terms, from their stated preferences. Whites' rank SP choices are consistent with a preference for neighborhoods with approximately 30%-50% Latino representation, and resistance to neighborhoods that are greater than 50% Latino. Whites state sharper resistance to Latino neighbors in the ideal SP vignettes, with lower choice probabilities beyond 40% Latino. After introducing all the non-racial controls, Whites' actual neighborhood choices

indicate a general unwillingness to live with Latinos, although with a wrinkle: Unlike the SP data, which show a mild preference for limited mixing with Latinos, the RH data suggest a monotonic decrease in choice probabilities as proportion Latino increases, albeit with higher choice probabilities in the 80-100% Latino range than in either SP scenario. However, statistical tests of SP-RH differences, in Table 2.8 Panel B, suggest no differences between rank SP and RH proportion Latino effects. These statistical tests do indicate differences between ideal SP data and RH data. Overall, then, Whites' preferences for living with Latinos appear to qualitatively align with their residential histories, and statistically align if we pay attention to the rank SP data.

In substantive and statistical terms, Whites come close to matching their preferences for Black neighbors. Figure 2.8 shows predicted probabilities across SP and RH scenarios as a function of neighborhood proportion Black. The rank SP data reveal Whites to be unlikely to choose neighborhoods as the proportion Black increases. The ideal SP predicted probabilities show an aversion to Black neighbors as well, but also a preference for some level of mixing—Whites are more likely to select 20% Black neighborhoods as ideal than they are to select 0% Black neighborhoods. This sharp preference for limited mixing with an out-group is common across ideal SP results for all groups. The predicted probabilities based on Whites' residential histories appear to align with either the rank SP or ideal SP predicted probabilities, depending on the non-racial controls employed. Prior to introducing controls for selection into migration, I find alignment between rank SP and RH predicted probabilities. Once I introduce all of the socioeconomic controls, Whites' RH choices suggest a preference for moving to neighborhoods that are 20% Black, and avoidance of neighborhoods that are more than 40% Black. However, statistical tests contained in Table 2.8 indicate that the rank SP-RH differences are not statistically significant, while the ideal SP-RH differences are. In general, I find qualitative agreement between Whites' stated preferences for Black neighbors and their real residential choices.

Whites come close to matching their preferences for neighborhood co-residence with other

Whites. Figure 2.9 displays predicted neighborhood choice probabilities as a function of proportion White. In the rank SP Data, Whites' choices were consistent with preferences for neighborhoods with greater than 50% White representation, and aversion to neighborhoods with low White representation. A similar pattern played out in the RH data, but with a stronger tendency to move to neighborhoods with higher White representation, peaking at approximately 80% White. The ideal SP curve reveals an even stronger in-group preference among Whites. Statistical tests of SP-RH differences in Table 2.8 Panel B suggest that I cannot reject the hypothesis of no difference between rank SP and RH. However, as suggested by the curves in Figure 2.9, I do find significant differences in Whites' ideal stated preferences, and their actual choices.

The above results suggest that Whites do achieve SP-RH matches when it comes to exposures to Latino, Black, and White neighbors. However, Whites do not appear to match their preferences when it comes to Asian neighbors. Both the Wald tests of ideal SP-RH and rank SP-RH differences in the effect of proportion Asian in Table 2.8 Panel B lead me to reject the null hypothesis of no SP-RH differences. Figure 2.10 shows that this result only appears after introducing controls for selection into migration. When Whites move, their subsequent residential choices bear only a weak, negative relationship with proportion Asian, despite stated aversion to neighborhoods that are more than 50% Asian in both SP reports. Thus, Whites in Los Angeles tend to move to neighborhoods with more Asian residents than they prefer.

2.5.4 Black Respondents

Across metropolitan contexts and surveys, Blacks express a desire to live in mixed neighborhoods (Farley et al. 1978, 1993; Charles 2006, 2000; Clark 2002; Krysan and Farley 2002; Krysan et al. 2009). Despite these preferences for mixing, across times and metro areas, Blacks tend to live in and move to predominately Black neighborhoods (Pais et al. 2012; South et al. 2008). This disconnect between preferences and neighborhood attainments does

not appear to hold in Los Angeles County. Table 2.7, Panel C presents Wald tests of sets of racial composition effects for Blacks. In the RH data (Panel B, column 1), the Wald tests indicate that Blacks' actual residential choices are unaffected by neighborhood racial composition save for a sensitivity to the presence of other Black residents ($p < 0.05$ for Black terms). However, in both SP scenarios (Panel C, Columns 1 and 2) Blacks are sensitive to the other racial composition components as well, except for only marginally significant sensitivity to proportion White in the rank SP scenarios ($p < 0.10$).

Blacks in Los Angeles County are a disadvantaged group, with lower levels of income and education than Whites and Asians, although generally higher than Latinos. Blacks also face some of the worst negative stereotypes. I predicted that they, like Latinos, would face barriers to matching their neighborhood circumstances to their racial composition preferences. Table 2.8, Panel C presents Wald tests of RH-SP discrepancies for Blacks. Column 1 of Panel C shows that Blacks' ideal stated preferences significantly deviate from their actual residential experiences for all racial composition components. In contrast, tests of differences between rank SP and RH (Column 2 of table 2.8) reveal no significant differences. A test of all racial composition variables simultaneously reveals only marginally significant rank SP-RH disparities ($p < 0.10$).

Because of Blacks' disadvantaged status, I expected that Blacks would be more likely to live with neighbors from other disadvantaged groups in Los Angeles, namely other Blacks and Latinos, in comparison to their preferences. Likewise, I expected they would have difficulty matching their preferences to live in neighborhoods with advantaged groups, i.e., Whites. Figures 2.11 through 2.14 demonstrate whether the statistically significant SP-RH discrepancies correspond, qualitatively, to the mismatches I hypothesized.

Blacks state a preference for and move to neighborhoods with moderate Latino representation. Figure 2.11 presents the relevant predicted probabilities. Blacks' rank SP choices are consistent with indifference to neighborhoods in the 0 to 40% Latino range, and aversion beyond 40% Latino. Blacks' RH choices suggest aversion to neighborhoods with high Latino

representation as well, with the probability of choice declining monotonically as a function of Latino representation. Once I introduce controls for selection into the migration stream, I find qualitative and statistical alignment between the rank SP and RH choices. The ideal SP choices deviate from the other two in that they feature narrow preferences for neighborhoods with some Latino representation, with peak choice probabilities in the range of 10 to 30% Latino representation. Blacks state a strong resistance to neighborhoods outside this range in the ideal SP data. If we go by the ideal SP data, Blacks do not match their preferences to their actual neighborhood choices, a result confirmed by the Wald tests in Table 2.8, Panel C.

Blacks' preferences for Black neighbors largely match their residential experiences. Figure 2.12 depicts neighborhood choice probabilities as a function of proportion Black. Blacks prefer neighborhoods that are roughly 50% Black, and resist choosing neighborhoods that have either very low or very high representation of Blacks, at least according to their rank SP responses. Blacks' RH choices suggest similar preferences, although with more resistance to all Black neighborhoods, and peak choice probabilities assigned to neighborhoods that are roughly 40% Black. The ideal SP results are similar to the RH results, but with a sharper peak predicted probability at 40% Black. Despite these qualitative differences, Wald tests of rank SP-RH and ideal SP-RH differences in the effect of proportion Black are not statistically significant. This all suggests that Blacks achieve an SP-RH match when it comes to neighborhood proportion Black.

Blacks in Los Angeles also do not appear thwarted in attaining their desired level of neighborhood co-residence with Whites. Indeed, the predicted probabilities in Figure 2.13 show that the rank SP and RH curves nearly match up, with Blacks preferring and tending to move to neighborhoods with 40% White representation. Again, the ideal SP curves depict an exaggerated preference for mixing, here near the 30% White composition level. However, differences between SP and RH in the effect of proportion White are not statistically significant, according to Table 2.8, Panel C. These results suggest that Blacks' residential moves

do not fall out of line with their preferences for neighborhood co-residence with Whites.

Finally, Blacks appear to move to neighborhoods with higher Asian representation than they would like, whether we compare RH data to ideal SP or rank SP data (Figure 2.14). However, these differences are only significant for the ideal SP-RH comparison, where Blacks once again have a sharp preferences for some mixing with the out-group, in the 10 to 30% Asian range, and resistance to neighborhoods outside this range. In the rank SP-RH case, both curves trace out a resistance on the part of Blacks to neighborhoods that are more than 40% Asian.

2.5.5 The Effect of Non-Racial Controls

In each of the previous graphs, I plot a number of RH predicted probability lines, corresponding to RH models that include different non-racial controls. Remarkably, once they are all layered in, these controls do little to affect the shape of the predicted probabilities. In all but a few cases, predicted probability lines for RH models that include only racial composition effects closely resemble those for RH models that include all non-racial controls in addition to racial composition. This implies that I can reject my hypothesis that I can account for SP-RH differences by controlling for non-racial factors intervening in processes of residential choices.

2.5.6 Putting RH-SP Mismatches in Context

The plotted predicted probabilities discussed above abstract from the extant residential possibilities in Los Angeles. The predicted probabilities are based on a neighborhood choice set that includes all possible racial compositions. However, a city whose neighborhoods span all possible racial mixes would be quite exceptional. While Los Angeles County is a diverse metropolitan area, in reality those seeking housing within the county's boundaries must choose among neighborhoods that encompass a limited subset of all possible racial compositions. Some racial composition configurations are in ample supply, and other racial

composition configurations are not observed. In particular, Los Angeles has a large number of mostly Latino and mostly White neighborhoods, and relatively few neighborhoods that are mostly Asian or Black. In addition, Los Angeles neighborhoods with higher White representation tend to have lower levels of Black and Latino representation, and slightly higher levels of Asian representation. Meanwhile, neighborhoods with high Black representation tend to have lower levels of Asian representation. In other words, when making migration decisions on the basis of race, residents of Los Angeles do so under constraints imposed by the availability of certain kinds of neighborhoods, and the bundling of particular neighborhood racial characteristics.

To address the relative supplies of neighborhoods in Los Angeles when interpreting model coefficients, I generate predicted probabilities based on the set of observed Los Angeles neighborhoods, rather than a hypothetical set of neighborhoods as in the previous discussion. These predicted probabilities indicate the degree to which SP assessed preferences differ from RH choices for a typical respondent facing a choice set of neighborhoods whose racial compositions match those in Los Angeles, but holding all other neighborhood characteristics equal.

I summarize the discrepancies between the RH and SP predicted probabilities for a Los Angeles County neighborhood choice set using indices of dissimilarity. These dissimilarities are given by:

$$D_{lm} = \frac{1}{2} \sum_{j=1}^N |p_{lj} - p_{mj}| \quad (2.7)$$

l and m distinguish coefficients derived from different data sources (e.g., RH, ideal SP, and rank SP), j indexes the N Los Angeles neighborhoods in the choice set, and p_{lj} and p_{mj} denote the predicted probabilities of choosing neighborhood j based on data source l and m , respectively. The values of this dissimilarity index range from 0 to 1, with 0 indicating a full

match between the predicted probabilities, and 1 indicating total mismatch. Alternately, the dissimilarity index can be interpreted as the proportion of neighborhood choices that would have to be changed in order to generate agreement between the two sets of predicted probabilities. I calculate pair-wise dissimilarities separately for the three racial groups previously discussed: Latinos, Whites, and Blacks. I calculate pair-wise dissimilarities between ideal SP and rank SP predicted probabilities, and a hypothetical “indifferent” housing seeker, who assigns the same choice probability to every neighborhood. I also produce dissimilarities based on two RH specifications, one that only includes racial composition terms, and the full model that includes controls for all previously discussed non-racial factors.

Table 2.9 provides evidence that Whites come closest to matching their stated preferences to their actual neighborhood circumstances. Whites have the lowest dissimilarities between RH predicted probabilities (Model 3A and 3B) and SP predicted probabilities (Models 1 and 2), whether I control for non-racial factors or not. Blacks come close to matching their rank SP assessed preferences, but only after accounting for non-racial factors. However, their residential histories are sharply discrepant with their ideal stated preferences. Latinos come closer than do Blacks to matching their ideal stated preferences and their neighborhood attainments. However, Latinos endure greater rank SP-RH dissimilarity than do Blacks once I control for non-racial factors. Latinos also experience greater SP-RH dissimilarities than Whites do.

I also use these predicted probabilities to predict the average neighborhood compositions experienced by typical Los Angeles Latinos, Blacks, and Whites when choosing from a Los Angeles neighborhood choice set. These predicted compositions are shown in Table 2.10. These predictions show whether the above dissimilarities are in the direction of over or under-exposure to certain groups, addressing my hypotheses about the likely over-exposure of Latinos and Blacks to other Latinos and Blacks, and low exposure of Whites to these groups.

The previously discussed ambiguous results for Latinos shine through in Table 2.10 Panel

A. Comparing predicted compositions based on ideal SP and RH predicted probabilities, I find that Latinos are actually under-exposed to Latino neighbors relative to their ideal preferences. Latinos would live in neighborhoods that are, on average, 75.1% Latino if they followed their ideal preferences, but only would live in 64.5 or 61.5% Latino neighborhoods based on RH model coefficients with and without non-racial controls, respectively. I find the opposite when comparing rank SP and RH. Using the rank SP assessment of preferences, I find that Latinos live with far more Latino neighbors than they wish (62.5-64.5% RH vs. 49% rank SP). This over-exposure to Latinos largely comes at the expense of under-exposure to Whites.

The predicted racial composition of Whites' neighborhoods, when faced with a Los Angeles County neighborhood choice set, largely aligns across rank SP and RH scenarios (Table 2.10 Panel B). Given a choice set of Los Angeles neighborhoods, Whites on average would live in neighborhoods that are approximately 55% White, 25% Latino, 15% Asian, and 5% Black when I employ either rank SP or RH coefficients in performing the calculations. There is some deviation between these results and the ideal SP results. Based on ideal SP coefficients, I would expect Whites to live in neighborhoods that are over 65% White. Whites seem to fall short of the high targets for White representation in their ideal neighborhoods. However, by any assessment, Whites prefer and attain residence in neighborhoods that are majority White.

The predicted racial composition of Blacks' neighborhoods roughly follows the previously discussed agreement between Blacks' rank SP preferences and their residential histories. Table 2.10 Panel C contains the relevant predictions. If Los Angeles Blacks followed the preferences expressed in the neighborhood ranking vignette, Blacks would live in neighborhoods that are roughly 36% Latino, 36% White, and 13% Black. Results from the RH models with full non-racial controls imply a similar racial composition, with slightly higher Latino representation and slightly lower White representation. In line with previously discussed ideal SP results for Blacks, if Blacks made choices based on their ideal preferences, their

neighborhoods would, on average, contain more White, Black, and Asian residents, and substantially fewer Latino neighbors. That is, Blacks’ ideal preferences would lead them to live in more racially heterogeneous neighborhoods than they actually live in. This result reflects the sharp preference for mixing expressed by Blacks in the ideal neighborhoods instrument.

2.6 Discussion and Conclusion

This study has estimated models of discrete neighborhood choice to characterize Los Angeles residents’ neighborhood racial composition preferences, both their stated preferences, and those preferences “revealed” in their residential histories. The conditional logistic regression models used here allowed me to combine two types of data, residential history and stated preference data, in a single unified model, which I then used to test explicit hypothesis concerning differences between stated preferences and residential histories across four racial composition dimensions: proportion Latino, proportion White, proportion Black, and proportion Asian. The modeling framework permitted the inclusion of controls for non-racial factors, at both the neighborhood and individual level, influencing neighborhood choices in residential histories. By comparing coefficients and predicted probabilities across stated preference and residential history portions of the unified model, and including non-racial controls in the RH portion of the statistical models, I evaluated the degree to which neighborhood attainments and preferences coincided across SP and RH observations, net of socioeconomic sorting factors.

Regardless of racial identification, residents of Los Angeles experience some mismatch between their stated racial composition preferences and the racial composition of their actual neighborhoods. However, some groups have more success matching their preferences than others. In particular, Whites come closest to matching their preferences to their neighborhood attainments. They both prefer to live in, and do live in, neighborhoods that are majority White. Latinos and Blacks face the greatest challenges in matching their residential

circumstances to their preferences, although with some ambiguities for both these groups. If I give credence to the rank SP data, then Blacks' preferences are more or less in alignment with their neighborhood circumstances. If I instead favor the ideal SP assessments, then Blacks' neighborhoods are more Latino and less heterogeneous than they wish. In contrast, the rank SP results for Latinos suggest that Latinos fail to match their neighborhood preferences, and tend to live in and move to neighborhoods with more Latinos and fewer Whites than they wish.

These results partly confirm a modified version of place stratification theory: Blacks and Latinos, two economically disadvantaged and negatively stereotyped groups in Los Angeles, encounter barriers to matching their neighborhood preferences. These barriers are not "explained" by accounting for the matching of individuals to neighborhoods based on income, education, home ownership status, and spatial proximity. This implies, although by no means proves, that Blacks and Latinos continue to face housing market discrimination, and provides support for place stratification theories.

That said, SP-RH discrepancies do not appear to be dramatic, even for Latinos and Blacks. It does not appear to be the case that any group's residential attainments run completely counter to their stated preferences. This is not to say that there are no individual group members whose experiences contradict their preferences. I have done little to account for heterogeneity in race-based preferences and attainments in this analysis. Future research should seriously consider whether there are some groups for which a large share of members experience substantial SP-RH disparities, despite overall agreement between SP and RH racial composition effects for "typical" group members.

Of course, the set of socioeconomic controls employed in the analysis do not exhaust the possibilities. I mostly have excluded wealth from this analysis. While I address home ownership as a factor influencing migration, I do not account for the value of the homes owned by different groups, nor do I account for possession of other liquid and illiquid assets that likely differ between groups and influence housing decisions. I have also not given

explicit consideration to housing costs, a factor correlated with, but distinct from levels of income and home ownership. In addition, housing prices and rents are determined at least partially through dynamic market mechanisms. Given a limited supply of housing within each neighborhood, more desirable neighborhoods will tend to have higher housing costs. This suggests a possible misspecification of the socioeconomic sorting mechanisms in the preceding models, which may lead to a misstatement of the effects of racial composition relative to non-racial factors. A modified model could account for competition among housing seekers for the most desirable neighborhoods (de Palma et al. 2007).

Another mechanism left out of this investigation is that of kinship. I would expect people to locate in neighborhoods that are spatially proximate to family members, especially close family members. Those family members will often share the same racial identifications as those making housing choices, and will likely live in neighborhoods that reflect the underlying patterns of racial segregation in the city. In this way, I may be overstating the role of racial composition in housing choices by ignoring the tendency of people to move to neighborhoods spatially proximate to their kin. Unfortunately, data on the residential locations of kin are difficult to come by. Even the Panel Study of Income Dynamics, with its multiple generations of residential history reports, incompletely tracks the residential locations of the kin whom we would expect to influence neighborhood choices.

The finding that Latinos are unable to attain the degree of co-residence with Whites that they prefer suggests that Latinos may face housing market discrimination. However, in addition to the potential misspecification of socioeconomic sorting, and the omission of kinship, the Los Angeles Latino population is heterogeneous in terms of nativity, duration in the United States, and national origins. It is possible that native born and long standing residents of the United States are better able to match attainments to preferences than those who arrived in the United States more recently. Further individual-level controls need to be employed to test this possibility.

The finding that Blacks managed to move to neighborhoods that matched their racial

composition preferences appears anomalous when compared with previous studies of Blacks' residential attainments, even in Los Angeles. This result could be attributed to four factors. First, this is the first study to use a discrete choice framework to compare race-based neighborhood preferences and attainments, and one of the few studies to explicitly account for the supply of destination neighborhoods, as well as the spatial proximity of these neighborhoods. Previous studies of Blacks in Los Angeles (e.g., Charles 2006, 2000) have not accounted for these factors. Second, Blacks represent a relatively small fraction of Angelenos, less than ten percent as of 2010. Neighborhoods that match Blacks' preferences for roughly 40% Black representation are rare: Less than 6% of neighborhoods fall in the 30% to 60% Black representation range. Only 1% of Los Angeles neighborhoods exceed 80% Black representation. While Blacks prefer neighborhoods with significant, but not overwhelming, shares of Black residents, there are many fewer extant Los Angeles neighborhoods on the majority Black side of this preference than there are on the minority Black side. In a sense, Blacks may not be forced into frequent compromises between their preference for mixed neighborhoods and the dearth of such neighborhoods, as may be the case in other metropolitan areas, like Detroit or Chicago. Third, it is possible that the practices of housing market discrimination have changed since the early 1990s, when the MCSUI data used for previous analyses were obtained. Finally, the sample of Black respondents in L.A.FANS is relatively small. While I report some statistically significant results for this group, the idiosyncrasies of the sample may render it poorly representative of the rest of the Los Angeles County Black population.

Overall, the discrete framework presented here represents an advance over previous attempts to judge the correspondence between stated preferences and actual neighborhood choices. Previous attempts have tended to privilege an absolute interpretation of the relationship between preferences and neighborhood attainments, implying that the neighborhoods that people say they prefer will correspond to real, plausible migration destinations. But this is not the case. People cannot simply choose any neighborhood they wish, but must choose from a limited set of available neighborhoods within their cities. It is in rela-

tion to the real set of possible neighborhoods that we must judge the relationship between individuals' migration decisions and their actual residential attainments. In this regard, the discrete choice analytical framework provides an important analytical microscope. In the next chapter, I provide a more in depth view of this framework.

Chapter 3

Discrete Choice Models for the Joint Analysis of Residential Preference and Residential History Data

This chapter develops an analytical framework for evaluating the degree of agreement between the neighborhood racial composition preferences people express, on one hand, and the actual neighborhood choices they make, on the other. This framework is needed to cast new light on debates about the relative roles of race-based residential preferences, socioeconomic inequalities, racial discrimination, and other factors, such as kinship networks, in generating and sustaining high degrees of racial residential segregation (Clark 1991; Galster 1988; Harris 1999; Krysan and Farley 2002; Krysan 2002). Without an appropriate analytical framework, these debates are untethered, with diverse methods brought to bear in ways that often ignore key components of the arguments under scrutiny.

One key question in these debates is whether people from different racial backgrounds enjoy the same amount of success in matching their residential preferences to their residential attainments. Scholars involved in these debates have harnessed two kinds of data to answer this question. Stated preference (SP) data catalog responses to hypothetical neighborhoods that vary in their racial composition. Respondents indicate in which neighborhoods they would feel comfortable, to which neighborhoods they would be willing to move, or to which neighborhoods they would most like to move. These data are intended to provide direct assessments of preferences. Residential history (RH) data track in which neighborhoods people

live, to which neighborhoods they move, and characteristics of individuals and their families that are pertinent to residential choices. Combining RH data with data on the existing, real neighborhoods in respondents' metro areas can reveal how race and racial composition influence migration. However, because of unobserved constraining factors in the housing market, including racial discrimination, these data may not correspond to preferences. An ideal analysis would combine both SP and RH data to evaluate the degree to which people are impeded in matching their preferences, thus providing an indirect assessment of the influence of unobserved constraining factors in inter-neighborhood migration.

I develop a discrete choice logistic regression framework for combining SP and RH data (Bruch and Mare 2012; Ben-Akiva and Lerman 1985; McFadden 1978, 1980). The discrete choice framework I develop differs from previous attempts to evaluate the correspondence between preferences and outcomes in three key ways. First, the framework accounts for differences in the “realms of the possible” housing seekers encounter when making real residential decisions and when stating their preferences. Typically, and by design, respondents do not select from the same sets of neighborhoods in the SP and the RH cases. Hypothetical neighborhoods in SP situations, and real neighborhoods in RH situations, often span incongruous ranges of racial compositions. Appropriate methods would take these incongruities into account. Second, the methods I use allow for concrete statistical tests of differences between racial composition preferences and the effects of racial composition on actual housing choices. Third, analytical results from the framework I develop here can be applied directly within simulation models of segregation (Schelling 1971; Bruch and Mare 2006; Clark and Fossett 2008). Simulations offer one approach to drawing out the implications of SP-RH mismatches, helping to resolve the question of whether segregation outcomes would differ if people actually got what they want.

The remainder of the chapter proceeds as follows. First, I review previous surveys that include SP or RH components, and past analyses of these SP and RH data. In the process, I highlight the advantages and disadvantages of these two data types. Second, I provide

a conceptual discussion of preferences, and the important conceptual and methodological problems that must be addressed by studies attempting to assess SP-RH agreement. Third, I lay out the discrete choice framework that I use to jointly model SP and RH outcomes. I discuss not only how this framework can be used to combine SP and RH data, but also how it can be used to combine data from multiple kinds of SP observations. Fourth, I discuss a random coefficients elaboration of this basic framework. These “mixed logit” models can be used to integrate together multiple observations per respondent and investigate preference heterogeneity. Fifth, I present three example cases. The first case combines multiple SP observations per respondent, where each of the observations is of the same “kind.” In the second case, I show how to combine SP responses to two differently phrased SP vignettes. Finally, I provide an example that combines SP and RH data, and then conclude.

3.1 Data for Assessing Racial Composition Influences on Neighborhood Choices: Stated Preferences and Residential History

Many previous studies have used stated preference and residential history data in isolation. These data types have contrasting strengths and weaknesses, but have not frequently been used together. The contrasting strengths and weaknesses of these data have been described at length elsewhere (Bruch and Mare 2012). Here I briefly recapitulate the review provided by Bruch and Mare. I build on their review by examining cases in which these two different data types have been combined and the methods used for these purposes.

3.1.1 Stated Preference Data

The SP approach uses choice experiments to assess neighborhood racial preferences (Louviere et al. 2000). Variants of this approach deploy survey based instruments or vignettes that solicit respondents’ opinions of hypothetical neighborhoods of varying racial compositions. Researchers have presented SP vignettes in a number of formats, including verbal descriptions

(Lewis et al. 2011), pictograms (Farley et al. 1978, 1993; Charles 2000), and videos (Krysan et al. 2009). A diverse set of methods have been brought to bear on these data. These methods include tabular and graphical approaches (Clark 1991, 2002, 2009; Farley et al. 1978, 1993), OLS regression (Charles 2000, 2006, 2007), multi-level modeling (Krysan et al. 2009), and logistic regression (Lewis et al. 2011). The chosen method varies with the outcome under examination and the experimental design. Whatever the method, studies employing these data have typically attempted to answer three questions. First, which neighborhood racial mixes do people find to be the most (un)desirable? Second, do members of racial groups differ, on average, in their desirability ratings? Third, what factors explain differences in preferences both within and between groups? These studies have largely shown that broadly defined Black, White, Latino, and Asian groups diverge in their assessments of neighborhoods according to racial composition. On average, members of all groups express a degree of own-group preference. But behind this own group preference, a racial hierarchy lurks, with Whites perceived as the most desirable neighbors, and Blacks the least desirable. Prejudices and stereotypes appear to play an important role: Those with strong negative stereotypes of a group also tend to prefer neighborhoods with low representation of that group.

3.1.2 Residential History Data

The “residential history” (RH) approach focuses on actual residential choices made by individuals and families in real housing markets. This approach typically assumes that neighborhoods are delineated by administrative boundaries, such as Census tract boundaries. Respondents either provide retrospective reports on the neighborhoods in which they have lived, as in the Los Angeles Family and Neighborhood Survey (L.A.FANS), or they provide these reports prospectively, as in the Panel Study of Income Dynamics (PSID). These data are linked to appropriate administrative or other data that describe the characteristics of neighborhoods, including their racial compositions. In analyzing RH data, individuals are assumed to have “chosen” neighborhoods whose boundaries contain their residential

addresses.

Models based on the resulting RH data incorporate measures of neighborhood racial composition and respondents' racial identifications to examine how these factors influence residential mobility and choices of destination neighborhoods. In the sociological literature, the assessment of RH outcomes has been dominated by two techniques: mobility models and spatial attainment models.

In studies using mobility models, neighborhood racial composition is included as a regressor to assess its effects on the out migration of individuals and families (Crowder 2000, 2001; Crowder et al. 2006; Crowder and South 2008; Crowder et al. 2012; South and Crowder 1997, 1998b; South et al. 2008). This dichotomous outcome is typically modeled using binomial logistic regression and its variants. This approach has been used to test the theory of "White Flight", which suggests that segregation emerged because Whites migrated to avoid Black neighbors during the middle of the twentieth century. Revised versions of this theory suggest that increasing segregation between Whites and expanding Latino and Asian populations results from Whites leaving neighborhoods to which Latinos and Asians have moved. Mobility studies reveal that Whites are prone to exiting neighborhoods with high minority concentrations, supporting at least one aspect of the White Flight theory.

Studies in the spatial attainment tradition examine neighborhood racial composition as an outcome in its own right, using linear regression to treat percent White in families' neighborhoods, or some other aspect of racial composition, as the dependent variable (Alba and Logan 1993; Sharkey 2012; South et al. 2005b, 2008, 2011). A few studies have revised the spatial attainment approach and treated neighborhoods categorically. These studies classify neighborhoods according to one or several neighborhood characteristics, including racial composition, and examine transitions into or between these neighborhood types (Alba and Logan 1991; South et al. 2005a; Crowder and South 2005; Crowder et al. 2006, 2012; Quillian 1999, 2002; South and Crowder 1998b). More recently, a number of authors have begun to use discrete choice models, the methods discussed in this chapter, to examine the

racial determinants of both out-migration and destination (Bayer and McMillan 2008; Bayer et al. 2007; Bruch 2014). Spatial attainment studies reveal that Whites are likely to move to neighborhoods with high White representation. Non-White groups appear to face barriers to entry into White neighborhoods, barriers that are not explained by lower, average levels of education, income, and wealth among non-whites.

3.1.3 Contrasting Strengths and Weaknesses of SP and RH Data

The SP and RH approaches have contrasting strengths and weaknesses. With the SP approach respondents express preferences unconstrained by other housing market forces. In the RH data, observed mobility, and thus the influence of racial composition on outcomes, is a product of preferences and constraints. If some sources of constraint or preference (e.g., financial limitations, wealth, housing market discrimination, a desire to live in a single family home, spatially patterned family obligations, etc.) cannot be fully accounted for in the RH data, then it becomes unclear if results pertaining to the effects of neighborhood racial composition on residential choice reveal racial preferences, or if race and racial composition variables capture unobserved preferences or constraining mechanisms. In this sense, the SP measures may provide purer depictions of preferences. But this supposed RH weakness also presents an opportunity when both SP and RH responses are available for the same respondents. If an analyst can account for differences between SP and RH by including variables that represent other constraints and preferences, this can be a powerful tool for understanding why residential outcomes differ across groups.

However, the SP approach's frequent exclusion of important non-racial aspects of neighborhoods, such as their socioeconomic composition or crime rates, can be a point of weakness. If individuals respond to neighborhood racial composition with other characteristics in mind, absent any true preference for racial mixing, this can lead to distorted depictions of how racial composition influences housing choices (Emerson et al. 2001; Harris 1999, 2001; Krysan et al.

2009). In some SP studies, researchers attempt to address this issue by providing respondents with more complex vignettes that identify nonracial aspects of neighborhoods. These attempts add to the cognitive load involved in comprehending hypothetical scenarios, potentially increasing response error. In addition, SP approaches may also be susceptible to social desirability bias when some respondents deliberately misstate their willingness to live in some types of neighborhoods. RH data, in contrast, measure individuals' responses to real neighborhoods that vary in racial composition and other characteristics. RH analyses that account for the full set of preference and constraint mechanisms can, theoretically, represent how race *truly* affects the housing selection process. However, as many of the relevant variables are either highly collinear or unobserved, it can be difficult to obtain appropriate estimates.

3.2 Combining SP and RH Data: A Critical Review

The complementary strengths of SP and RH data suggest that, where possible, they should be combined. In fact, to make explicit, statistical, and analytical judgements about whether people in the population tend to get what they want in the housing market, or whether they fail to match their preferences, studies *must* combine both types of data.

Few data sets have collected SP and RH data together, with the Multi-City Study of Urban Inequality (MCSUI) and the Los Angeles Family and Neighborhood Survey (L.A.FANS) standing as notable exceptions. Previous attempts to assess the SP-RH relationship, often with MCSUI data, have used tabular techniques or linear regression. These approaches, however, fail to address key conceptual issues, including a distinction between absolute and relative preferences, and the related issue of neighborhood choice sets. These approaches also neglect neighborhood variation along multiple, correlated dimensions; both racial/ethnic and socioeconomic. Below I review the results of previous attempts to compare SP and RH data, and then discuss the conceptual problems these studies raise. I then propose an analytical framework that attempts to resolve these issues.

3.2.1 Previous Work Combining SP and RH Data

Early approaches to comparing preferences to attainments used descriptive, tabular techniques. Clark (1992) maps out, in sparsely populated contingency tables, the association between stated, racial composition preferences and the racial composition of actual neighborhood attainments. Using data from a telephone survey of Los Angeles residents, Clark reports that all groups show a tendency to prefer own group representation in their neighborhoods, although with greater preferences for mixing among Blacks. At the same time, some groups are better able than others to realize their preferences in the housing market, with Latinos and Asians seemingly less successful in realizing their preferences relative to Whites.

Other attempts to compare revealed and stated preferences have used locational attainment models (Logan et al. 1996; Alba and Logan 1993). This approach regresses percent White in respondents' actual neighborhoods on the characteristics of respondents and their families, including income, education, housing tenure, wealth, and, importantly, neighborhood racial composition preferences (Adelman 2005; Freeman 2000). Proceeding from the assumption that preferences precede neighborhood outcomes, these analyses find that, across groups, those with greater preferences for White neighbors live in neighborhoods with more Whites. But, consistent with much previous research on residential attainment, Blacks are disadvantaged relative to Whites, Latinos, and Asians in gaining access to White neighborhoods, even when controlling for preferences and socioeconomic attainments

Unmodified OLS approaches based on cross-sectional data ignore the potentially reciprocal relationship between preferences and neighborhood outcomes. Not only might those who prefer White neighbors live in Whiter neighborhoods, but also those who live with White neighbors may come to prefer Whiter neighborhoods. Other studies in the spatial attainment paradigm deal with this potential endogeneity by estimating simultaneous equations models, treating preferences and neighborhood racial composition as simultaneous outcomes (Charles 2006; Ihlanfeldt and Scafidi 2002, 2004). For example, Charles (2006) models per-

centages Asian, Black, and Latino in individuals' preferred and actual neighborhoods as mutually determined outcomes. Differences between racial and ethnic groups' neighborhood attainments persist even when accounting for their stated preferences. In particular, Blacks and Latinos appear to be disadvantaged in attaining residence in neighborhoods matching their preferences, and are more likely to live in Black or Latino neighborhoods regardless of their preferences to do so, or their socioeconomic characteristics.

3.2.2 The Importance of Choice Sets

Both the tabular and spatial attainment approaches, while providing useful descriptions of the state of preferences or residential inequalities in particular metro areas, are ill-suited to comparing SP and RH data. Importantly, previous approaches to comparing SP and RH data glaze over the conceptual issues related to the intrinsically relative nature of preferences. Expressing a preference, whether in the SP or RH case, involves the ordering of sets of alternatives or outcomes according to their perceived qualitative and quantitative attributes. The expression of preferences is intrinsically bound up with the opportunities that people have to choose neighborhoods. Neighborhoods are typically conceptualized as discrete units that are presented as alternatives in a "choice set". The choice set contains all the neighborhoods to which a person may move, or for which each person may express a preference. Because neighborhoods can differ along many different dimensions, and in many gradations, the finite and limited size of this choice set typically will omit many logically plausible neighborhoods. Setting aside constraining factors in the RH case, an analyst can judge individuals' preferences for a neighborhood attribute only to the extent that the available neighborhood alternatives, contained in the choice set, perceptibly vary in these attributes, and to the extent that neighborhood choices/moves vary across alternatives according to these attributes.

Because people can express preferences only relative to a set of available alternatives, analysts must attend to the possibility that SP-RH differences crop up because of differences

in SP and RH choice sets, rather than because of differences in the ways that people order these neighborhood alternatives according to racial composition. A failure to account for choice set differences across SP and RH data can lead to exaggerations or under-statements about the degree of mismatch. That is, an analyst must consider the balance between preferences and opportunities.

In RH data, choice sets are made up of the full set, or sub-sets, of extant neighborhoods in the metropolitan areas where respondents reside. This poses two supply constraints. First people cannot move to neighborhoods that do not exist. Some racial mixes are not featured at all among a city's neighborhoods. In Los Angeles, for example, there are no tracts that are 40% White and 60% Black, nor are there any neighborhoods that are 50% Black and 50% Asian. Second, even among existing neighborhood racial mixes, some racial mixes occur more frequently than others. In Los Angeles, many neighborhoods feature an 80% Latino and 10% White ethnic mix, but there are very few neighborhoods that are 40% Latino and 20% White. These supply constraints together limit the realm of possibility for residential outcomes.

In SP data, survey designers determine the universe of possible neighborhoods that respondents choose from. Researchers decide how many and what neighborhoods to present to each respondent, and respondents can choose only neighborhoods within this choice set. In the MCSUI studies, for example, respondents were presented with neighborhoods that varied systematically in only two racial composition dimensions at a time: a respondent's own group was treated as the reference group, and then respondents were asked to rate neighborhoods as either the proportion Asian, Black, Latino, or White increased at the expense of the respondent's own group. The resulting neighborhood choice set often entirely excludes racial mixes that occur frequently in respondents' cities, or implicitly over-represents certain racial mixes. For example, including one 80% White, 20% Latino neighborhood and one 50% White, 50% Latino neighborhood in the choice set implies that these neighborhoods are in equal supply, when the actual, metro-area supply of housing units in those neighborhoods

may be quite unbalanced.

SP-RH choice set differences pose acute problems for using contingency tables to compare SP and RH data. Some cells may be empty because those cells correspond to neighborhoods in very low supply, either in the SP or RH data. People may appear to match their preferences simply because of an over-supply of a given type of neighborhood, or may be prevented from doing so because of an under-supply of a given type of neighborhood. It would be erroneous to read this as an indication that people, in general, are successful or fail in matching their preferences.

Choice set differences also pose problems for spatial attainment/linear regression approaches to analyzing SP-RH correspondence. Analysts make implicit assumptions about the available neighborhood alternatives via the normal distributional assumptions for error terms in the linear regression models. In many cases, in fact, by taking a percentage or proportion as the dependent variable, analysts implicitly assume that some patently impossible racial compositions are, in fact, plausible. Such is the case when estimated variances imply significant likelihoods of living in “neighborhoods” that fall below 0% or above 100% representation for a given group. Similar, but perhaps more subtle, issues occur when distributional assumptions encompass theoretically plausible racial compositions that are, nonetheless, observed in few or no existing neighborhoods in the SP or RH choice set.

3.2.3 Multi-Dimensionality in Neighborhood Choice

Both tabular and spatial attainment analyses of neighborhood outcomes also fail to account for the fact that neighborhoods vary across several dimensions. Oftentimes, these dimensions are correlated, either mechanically or as a result of social processes. In SP data where racial composition is the focus, neighborhood racial composition components (i.e., proportion Black, proportion Latino, etc.) are mechanically correlated. In the RH data, not only are the racial composition components mechanically correlated, but also they may be

correlated as a result of the social processes under study. In addition, racial composition may be correlated with other neighborhood characteristics, such as levels of poverty or affluence, that are important determinants of housing market choices.

Racial composition variables are mechanically correlated in SP and RH data. This occurs because, once the population is categorized into mutually exclusive groups, the sum of their proportions must come to one. This typically induces a negative correlation between these components, even in the SP case. Furthermore, in places like Los Angeles, with multiple racial and ethnic groups present in substantial proportions, the treatment of a single racial composition term, like percentage Black, as an outcome is unrealistic. Neighborhood racial compositions vary non-trivially across three or more dimensions in Los Angeles, and increasingly so in the rest of the United States. If respondents tend to move to neighborhoods with higher Black representation, this could indicate a preference to avoid White neighbors, but could also imply neutral feelings about Whites accompanied by a preference to avoid other non-White groups. Both the tabular and locational attainment approaches ignore how neighborhood racial ethnic compositions are interdependent, and that individuals negotiate racial preferences in racially and ethnically variegated contexts.

Social processes of residential segregation along non-racial dimensions, such as housing prices, income, and school quality, lead to correlations between racial composition and other factors that influence neighborhood choices in RH data. A cursory examination of Census data, for example, shows that neighborhoods with high levels of income also tend to have higher proportions of White residents. An analysis of exposure to White neighbors may misrepresent a preference for more affluent neighborhoods as a preference for White neighbors. The spatial attainment literature has vaguely addressed this problem by frequently estimating two neighborhood outcome models: one for exposure to Whites, another for exposure to poverty, or neighborhood median income. This does not resolve the issue. Such approaches continue to treat these racial composition and socioeconomic outcomes as independent.

3.2.4 Micro-Macro Linkages

Finally, contingency table and linear regression approaches are difficult to extend to simulation models of neighborhood mobility, like Schelling’s model of residential segregation (Schelling 1971; Fossett 2006a; Bruch and Mare 2006). This is because linear regression and contingency table approaches do not treat neighborhoods themselves as the outcomes, and instead treat attributes of neighborhoods as outcomes. OLS regression and its variants treat neighborhoods as univariate outcomes, like proportion Black. Contingency table approaches treat neighborhoods as representatives of categories, like high affluent White neighborhoods, high poverty Black neighborhoods, and so on. In both cases, it becomes unclear how to translate model results into the choices of hypothetical agents making choices in a simulated housing market. This can preclude or obscure an understanding of the macro-level implications of any between-group differences in processes of neighborhood attainment. To grasp macro-level consequences, models that complement simulation approaches should consider the fact that individuals move between neighborhoods and housing units, not between percentages Black or White or Latino, or between neighborhood categories.

In sum, previous approaches neglect the importance of choice sets, ignore the variation of neighborhoods along multiple, inter-related racial composition and non-racial dimensions, and are difficult to apply in simulations models that elucidate the link between micro behaviors and macro patterns. In the following sections I develop models that address all of these issues.

3.3 Discrete Choice Models for Combining SP and RH Data

Discrete choice models address the methodological and conceptual problems ignored by conventional methods for making SP-RH comparisons. Discrete choice models explicitly account for differences in choice sets to assess how respondents in SP and RH data make

relative trade-offs between neighborhood attributes. The approach permits the inclusion of multiple, potentially correlated neighborhood-level regressors as determinants of neighborhood outcomes, and results from discrete choice models can be applied to simulation models of residential attainment. Below I describe the basic discrete choice approach. I then consider the methodological question of how to combine SP and RH data using discrete choice methods.

3.3.1 Discrete Choice: Conditional Logistic Regression

The models I use are elaborations of conditional logistic regression, an approach that has substantial precedent in the econometric literature (McFadden 1978; Ben-Akiva and Lerman 1985; Train 1986; McFadden 2001) and has recently been reintroduced to the sociological literature (Bruch and Mare 2012). In the simplest discrete choice model, individuals, indexed by i , are presented with a *choice set*, C_{it} , which may vary across individuals and decision instances, indexed by t . The choice set contains the set of possible neighborhood *alternatives*, indexed by j , to which an individual can move. A person assigns a utility, U_{ijt} , to each neighborhood in his choice set in a given choice instance according to a presumed function of the characteristics of each neighborhood, j . Finally, agents choose the neighborhood alternative that provides the highest utility.

The discrete choice approach assumes that neighborhood utilities, U_{ijt} , can be decomposed into an observed and unobserved part. The observed characteristics are captured by V_{ijt} , while the unobserved characteristics are captured in the random disturbance, ϵ_{ijt} .

$$U_{ijt} = V_{ijt} + \epsilon_{ijt} \tag{3.1}$$

An analyst parameterizes the “observed” part of the utility according to the observed characteristics of neighborhoods that are thought to influence neighborhood choices.

$$V_{ijt} = \beta_{it}X_{jt} + \zeta_{it}Z_{ijt} \tag{3.2}$$

In the above formulation, X_{jt} is a column vector of observed, objectively evaluated neighborhood racial attributes, like percent Black or percent Latino. β_{it} is a row vector of coefficients which denote the weight given to the racial attributes of neighborhoods in the decision process. The i subscript for β indicates that the effects could differ across individuals, and the t subscript indicates the effects could differ across choice instances. Conceptually, the β 's are the racial composition preferences. They determine the contributions of the racial composition components to the underlying utility, and thus determine the relative positions of neighborhoods in the (assumed) utility hierarchy that underpins the model.

The second piece of observed utility, $\zeta_{it}Z_{ijt}$ represents the contribution of non-racial factors to the overall utility. Z_{ijt} is a column vector of objectively or subjectively evaluated non-racial neighborhood characteristics. Objectively evaluated non-racial neighborhood attributes are fixed for a given alternative, regardless of the respondent making the choice. These could include neighborhood poverty, median income, or age composition. Subjectively evaluated neighborhood attributes vary by respondent. These could include the distance of a neighborhood from an individual's place of work, whether or not a close family member is residing in the neighborhood, or whether the individual already lives in the neighborhood. ζ_{it} is a row vector of coefficients which indicates the influence of these non-racial aspects of neighborhoods in residential choices. The subscripts indicate these coefficients could differ across time and across agents. This means that the effects of both objective and subjectively evaluated neighborhood attributes could vary depending on the respondent's own characteristics, or the historical moment in which a choice is made.

By assuming the disturbances, ϵ_{ijt} , are uncorrelated with the observed variables and that they are drawn independently from a standard Gumbel distribution, the probability that individual i chooses neighborhood j in scenario t is given by:

$$P_{ijt} = \frac{\exp(\beta_{it}X_{jt} + \zeta_{it}Z_{ijt})}{\sum_{k \in C_{it}} \exp(\beta_{it}X_{kt} + \zeta_{it}Z_{ikt})} \quad (3.3)$$

Given a sample of N individuals, indexed by i , who each make a sequence of T choices, $\{j_{i1}, j_{i2}, \dots, j_{iT}\}$, from known choice sets C_{it} , a likelihood function can be constructed based on the above expression for choice probabilities, if I assume that the unobserved disturbances are uncorrelated across individuals and across choice instances:

$$\mathcal{L} = \prod_{i=1}^N \prod_{t=1}^T \frac{\sum_{j \in C_{it}} y_{ijt} \exp(\beta_{it} X_{jt} + \zeta_{it} Z_{ijt})}{\sum_{k \in C_{it}} \exp(\beta_{it} X_{kt} + \zeta_{it} Z_{ikt})} \quad (3.4)$$

Here I have introduced the outcome variable, y_{ijt} . This is a dummy variable that takes on the value 1 if neighborhood j was chosen by person i at time t , and is 0 for all other neighborhoods in person i 's choice set at time t . For the first choice made by person i , $y = 1$ if $j = j_{i1}$, and is zero for all other neighborhoods. For person i 's second choice, $y = 1$ if $j = j_{i2}$ and is zero for all other neighborhoods, and so on.

Traditional maximum likelihood methods can be used to estimate β_{it} and ζ_{it} . The estimated β_{it} indicates which neighborhood racial attributes increase or decrease the likelihood of choosing a neighborhood. Note that direct effects of individual characteristics drop out of the probability expressions and the likelihood, meaning that they are not estimated or estimable. Instead, individual characteristics enter the model only in interaction with neighborhood characteristics.

To take an example, X_{jt} could include a variable describing the proportion of neighborhood j 's residents who identified as White in census data,. The corresponding β_{it} would indicate the average "preference" for this characteristic among respondents. A positive coefficient would imply that respondents prefer the presence of Whites in their neighborhoods. A negative coefficient would indicate that respondents are averse to White neighbors. We could examine preference differences between broad racial groups by including an interaction of neighborhood racial composition terms with dummies indicating racial identification, or by estimating choice models separately by respondent race.

3.3.2 Conditional Logit Models for Combining SP and RH Data

Combining data from SP and RH sources is a natural extension of the discrete choice model, again with substantial precedent in the econometrics literature (Ben-Akiva et al. 1994; Brownstone et al. 2000; Hensher and Bradley 1993; Hensher et al. 2008; Morikawa 1989, 1994). The key maneuver involves noting that the utility functions, in principle, might differ between SP and RH data types. This could be for two main reasons. First, the set of regressors available in the SP and RH data might differ, often because the SP vignettes include only a subset of variables that can effect choices in the “real world” RH data. Second, even for variables that are common across SP and RH data, we might expect effects of these variables to differ across data types. This is the case for neighborhood racial composition, where processes of racial discrimination, or social network influenced migration, can push RH and SP racial composition effects out of line. The possibility of different SP and RH effects is already encoded in Equation 3.3, as β_{it} can differ across respondents and choice situations. However, below I make the division between SP and RH more explicit.

Consider a respondent, i , who has provided both SP and RH data. In the SP data, the respondent is presented with choice sets C_{is} across choice situations $s = 1, 2 \dots S$. I write the utility for each alternative, k , in each SP instance s as:

$$\begin{aligned} U_{iks}^{SP} &= V_{iks}^{SP} + \epsilon_{iks}^{SP} \\ V_{iks}^{SP} &= \beta_{is}^{SP} X_{iks}^{SP} + \zeta_{is}^{SP} Z_{iks}^{SP} \end{aligned} \tag{3.5}$$

Where V_{iks}^{SP} is the observed portion of utility and ϵ_{iks}^{SP} is an unobserved error term. X_{iks}^{SP} is a column vector of racial composition components and Z_{iks}^{SP} is a column vector of non-racial components for alternative k . I use the SP superscript to indicate that these sets of explanatory variables may differ from those in RH data.

In RH data, I index the T instances of residential choices by $t = 1, 2 \dots T$. In each instance, t , a respondent faces a choice set C_{it} . I write the utility for each neighborhood, j ,

in this choice set as:

$$\begin{aligned} U_{ijt}^{RH} &= V_{ijt}^{RH} + \epsilon_{ijt}^{RH} \\ V_{ijt}^{RH} &= \beta_{it}^{RH} X_{ijt}^{RH} + \zeta_{it}^{RH} Z_{ijt}^{RH} \end{aligned} \tag{3.6}$$

The RH super-script distinguishes the racial composition variables, X_{ijt}^{RH} , and non-racial variables, Z_{ijt}^{RH} , from those in the SP data. In many attempts to elicit racial composition preferences, Z_{iks}^{SP} is entirely excluded from the data set, and so must be excluded from the analysis.

If the racial composition variables included in the RH and SP data are the same, the RH and SP superscripts on X can be dropped. This shifts attention to the coefficients. The β 's are individual specific vectors of racial composition effects that potentially vary across SP and RH cases. There are separate β 's for RH and SP scenarios, which may differ for reasons related to social desirability bias, the artificiality of the SP choice scenarios, or because of unobserved constraints in RH choice scenarios. We can write the β 's as:

$$\begin{aligned} \beta_{it}^{RH} &= \beta_{i0}^{RH} + \gamma^{RH} G_{it} \\ \beta_{is}^{SP} &= \beta_{i0}^{SP} + \gamma^{SP} G_{is} \end{aligned} \tag{3.7}$$

β_{i0} represent represent unobserved, person-specific preference intercepts that may differ across SP and RH data. For a given choice scenario, these preferences are modified by an individual's observed, exogenous, and potentially time varying attributes, G_{is} and G_{it} . G could include observed factors such as a respondent's race, marital status, or age. These could be fixed within respondents, or time varying. G could also contain information about the choice situation itself. I assume that the set of factors included in G is the same for SP and RH scenarios. However, these individual-level factors may have different effects, γ^{SP} and γ^{RH} . In more complicated models, one can use β_{i0} to account for unobserved, heterogeneous preferences. RH and SP models must be estimated jointly to account for the unobserved, person-specific initial preferences represented by β_{i0} , which are potentially influencing all

SP and RH choices. I discuss these more complicated models in more detail in subsequent sections.

However, one can make a single simplifying assumption to arrive at more tractable models. The assumption is to ignore potential unobserved heterogeneity in β_{i0}^{RH} and β_{i0}^{SP} , setting these coefficients equal to fixed constants, β^{RH} and β^{SP} , respectively. This yields:

$$\begin{aligned}\beta_{it}^{RH} &= \beta^{RH} + \gamma^{RH}G_{it} \\ \beta_{is}^{SP} &= \beta^{SP} + \gamma^{SP}G_{is}\end{aligned}\tag{3.8}$$

This formulation is only subtly different from Equation 3.7, but the absence of the i subscript encodes a critical assumption that all heterogeneity is captured by the coefficients, γ^{RH} and γ^{SP} along with the exogeneous, individual and choice situation specific variables G_{it} and G_{is} . Estimates of these coefficients will result in separate racial composition coefficients for SP and RH cases.

It can also be helpful to re-parameterize the above to obtain estimates of SP coefficients as deviations from the RH case (or vice versa). I set $\beta^{RH} = \beta$, a constant, baseline effect in the RH case, and re-express β^{SP} as a residual deviation from this baseline, $\beta^{SP} = \beta + \Delta\beta^{SP}$. I also assume $\gamma^{RH} = \gamma$, and re-express γ^{SP} as a deviation from the (baseline) RH case: $\gamma^{SP} = \gamma + \Delta\gamma^{SP}$. These manipulations yield:

$$\begin{aligned}\beta_{it}^{RH} &= \beta + \gamma G_{it} \\ \beta_{is}^{SP} &= \beta + \Delta\beta^{SP} + \gamma G_{is} + \Delta\gamma^{SP}G_{is}\end{aligned}\tag{3.9}$$

Given a sequence of chosen RH and SP alternatives for the i^{th} respondent, $\mathbf{c}_i = \{\{h_{i0}, h_{i1}, \dots, h_{iT}\}, \{p_{i1}, p_{i2}, \dots, p_{iT}\}\}$, with h indexing this set of chosen RH alternatives, and p indexing the chosen SP alternatives, the likelihood for the full sample of N respondents is given by:

$$\mathcal{L} = \prod_{i=1}^N \prod_{t=1}^T \left[\frac{\sum_{h \in C_{it}} y_{iht} \exp((\beta + \gamma G_{it})X_{iht} + \zeta Z_{iht})}{\sum_{j \in C_{it}} \exp((\beta + \gamma G_{it})X_{ijt} + \zeta Z_{ijt})} \right] \times \prod_{s=1}^S \left[\frac{\sum_{p \in C_{is}} y_{ips} \exp((\beta + \Delta\beta^{SP} + \gamma G_{is} + \Delta\gamma^{SP} G_{is})X_{ips})}{\sum_{k \in C_{is}} \exp((\beta + \Delta\beta^{SP} + \gamma G_{is} + \Delta\gamma^{SP} G_{is})X_{iks})} \right] \quad (3.10)$$

Where y_{iht} and y_{ips} are dummy outcome variables denoting chosen neighborhoods in RH and SP scenarios, respectively.

This likelihood can be maximized with respect to the parameters using conventional maximum likelihood approaches. This is the joint SP-RH conditional logistic regression model. A number of software packages are capable of estimating this model, including Stata, R, LIMDEP, biogeme, and SAS.

In this formulation, it is valid to estimate separate models for the SP and RH data. In practice, however, it is sensible and easy enough to estimate the models jointly. This allows for straightforward estimation of constrained models. For example, we might be particularly interested in the case in which $\Delta\beta^{SP} = 0$ and $\Delta\gamma^{SP} = 0$, i.e., a model in which there is no difference in racial composition effects between SP and RH scenarios. Joint estimation also enables easy post-estimation hypothesis testing.¹

3.3.3 Combining Data from Multiple SP Sources

The same procedures for combining SP and RH data can be used to combine data from multiple SP sources. This approach may be appealing when stated preferences are assessed using multiple instruments or questions. In MCSUI, respondents expressed their preferences for researcher generated neighborhoods in two ways: respondents indicated which neighborhoods they found most attractive, and into which neighborhoods they would be willing to

¹Note that one can also estimate the likelihood using the specification implied by Equation 3.8. The approaches are statistically equivalent in the conditional logistic case, the only difference is in how the coefficients are interpreted. In a model specified as in Equation 3.8, the racial composition effects would be interpreted as separate, main effects in the SP and RH cases, rather than as a main RH effect, with the SP effect treated as a deviation from the RH effect.

move. In addition, the Los Angeles version of the survey asked respondents to identify their ideal neighborhoods’ racial compositions by filling out blank neighborhood cards with the “ideal” number of Asian, Black, Latino, and White neighbors. In L.A.FANS, respondents ranked five randomly generated neighborhoods, varying in the proportions Asian, Black, Latino, and White, according to where they would most like to live. L.A.FANS also asked respondents to complete an ideal neighborhood exercise. A single discrete choice model could be used to combine all these SP responses.

These different SP questions and vignettes might lead respondents to express different preferences. If different responses, and hence different racial composition effects, are expected for different SP tasks and questions, then it is fitting to adapt the above techniques for SP-RH comparisons for SP-SP comparisons. Even for a single question that elicits a ranking of neighborhoods, there could be differences in racial composition effects across ranks. The calculus people employ when selecting the first ranked neighborhood might differ from that used to rank the fourth ranked neighborhood over the fifth ranked neighborhood.

One need only reframe the RH portion of the model described above as a second type of SP data to test for differences across SP approaches. Procedures like this have been used to compare the effects of alternative specific covariates in other contexts, most notably in evaluating preferences for transportation modes (Ben-Akiva et al. 1991).

3.3.4 Accounting for Multiple Observations within Respondents: Mixed Logit

Ideally, SP and RH responses would come from the same set of respondents. This is the case for L.A.FANS, which I use in Chapter 2. In L.A.FANS, respondents expressed their preferences in two different SP vignettes, and also provided multiple years of residential history. Random coefficient, “mixed logit” models can leverage the panel nature of these data, and multiple SP and RH responses obtained for each respondent, to provide more realistic predictions, and to describe the degree of preference heterogeneity across the population (Hensher and Greene 2003; Revelt and Train 1998; McFadden and Train 2000). Mixed logit

models can also be used to relax the assumption that observations are independent within respondents.

To formulate the mixed logit model, I recall the initial expressions for utility in the SP and RH cases in Equation 3.7, prior to making the simplifying assumptions to make the model tractable in Equation 3.8. I rewrite the SP and RH utilities in Equation 3.7 in terms of the regressors, X , and coefficients, β , dropping the the non-racial terms, Z , for the sake of simplicity:

$$\begin{aligned} U_{ijt}^{RH} &= \beta_{it}^{RH} X_{ijt}^{RH} + \epsilon_{ijt}^{RH} \\ U_{iks}^{SP} &= \beta_{is}^{SP} X_{iks}^{SP} + \epsilon_{iks}^{SP} \end{aligned} \tag{3.11}$$

Prior to simplification, I expressed the β s as:

$$\begin{aligned} \beta_{it}^{RH} &= \beta_{i0}^{RH} + \gamma^{RH} G_{it} \\ \beta_{is}^{SP} &= \beta_{i0}^{SP} + \gamma^{SP} G_{is} \end{aligned} \tag{3.12}$$

In the previous discussion, I ignored unobserved heterogeneity by setting β_{i0}^{RH} and β_{i0}^{SP} to constants. This step generates models that treat multiple observations from the same respondents as independent. Clearly this is unrealistic: someone who has a strong preference for White neighbors will probably act on this preference with some consistency across time in RH data and across vignettes in SP data. And we might expect those with the strongest preferences for White neighbors in the SP data to act on those preferences in the RH data. Methods of modeling heterogeneity and explicitly acknowledging the interdependence of observations within respondents relax this restrictive independence assumption. In so doing, the models can lead to better behavioral predictions, and can help to answer more subtle questions about the distribution of preferences across the population.

The mixed logit model deals with unobserved heterogeneity by assuming a continuous mixing distribution for β_{i0}^{RH} and β_{i0}^{SP} . The mixed logit model has become an important tool

in the arsenal of discrete choice modelers, and substantial work has been done to apply these models across a range of cases (Hensher and Greene 2003; Revelt and Train 1998; McFadden and Train 2000). Several studies have explored the use of mixed logit for jointly modeling preference and observational choice data (Brownstone et al. 2000; Bhat and Castelar 2002). However, most of this work has assumed that effects in stated preference and observational choice portions of discrete choice models should coincide. Below, I outline a mixed logit framework that allows for effects to differ across SP and RH portions of the model.

Implementation of mixed logit models requires three steps in addition to those taken with the traditional conditional logit. First, one must specify the mixing distribution for the individual-specific effects. Second, a revised, unconditional likelihood function is formed to account for the dependence between observations. Finally, the new likelihood function must be estimated using modified maximum likelihood techniques.

Many different mixing distributions can be used, but in the racial composition case, the multi-variate normal seems most appropriate. This is because of the likely inter-related nature of racial composition preferences. It is not difficult to imagine, for example, that White respondents who prefer to avoid neighborhoods with high Black representation may also prefer to avoid neighborhoods with high Latino representation. The multi-variate normal distribution readily accepts preferences that are correlated across racial composition components. A multi-variate normal mixing distribution can also better represent the potential divide between SP and RH racial composition effects, allowing for the preferences in these cases to be correlated within respondents, but not perfectly so.

Thus I assume that β_{i0}^{RH} and β_{i0}^{SP} are drawn from a multi-variate normal distribution. In matrix notation:

$$\beta_{i0} \sim \mathcal{N}(\beta, \Sigma) \tag{3.13}$$

β_{i0} concatenates the individual specific coefficient vectors for the RH and SP data.

$$\beta_{i0} = \begin{bmatrix} \beta_{i0}^{RH} \\ \beta_{i0}^{SP} \end{bmatrix} \quad (3.14)$$

β is a vector that contains the population mean racial preference coefficients for the RH and SP scenarios.

$$\beta = \begin{bmatrix} \beta^{RH} \\ \beta^{SP} \end{bmatrix} \quad (3.15)$$

The variance-covariance matrix, Σ , contains sub-matrices that indicate variation and covariation of racial composition effects within SP and RH scenarios, and the correspondence of these effects across scenarios.

$$\Sigma = \begin{bmatrix} \Sigma^{RH} & \Sigma^{RH,SP} \\ \Sigma^{SP,RH} & \Sigma^{SP} \end{bmatrix} \quad (3.16)$$

Given a sequence of RH and SP choices, $\mathbf{c}_i$, for each individual, the likelihood for each respondent, conditional on β_{i0} , is given by

$$L_{i\mathbf{c}_i}(\beta_{i0}) = \prod_{t=1}^T \left[\frac{\sum_{h \in C_{it}} y_{iht} \exp(\beta_{i0}^{RH} X_{iht})}{\sum_j \exp(\beta_{i0}^{RH} X_{ijt})} \right] \times \prod_{s=1}^S \left[\frac{\sum_{p \in C_{is}} y_{ips} \exp(\beta_{i0}^{SP} X_{ips})}{\sum_k \exp(\beta_{i0}^{SP} X_{iks})} \right] \quad (3.17)$$

Where, again, y_{iht} is a dummy variable that takes the value 1 if person i moved to neighborhood h at time t in the RH data, and y_{ips} is a dummy variables that takes the value 1 if person i chooses neighborhood p in instance s in the SP data.

The unconditional likelihood is given by:

$$\mathcal{L} = \prod_{i=1}^N \int L_{i\mathbf{c}_i}(\beta_{i0}) f(\beta_{i0}|\beta, \Sigma) d\beta_{i0} \quad (3.18)$$

Where f is the multivariate normal density, and the integral is performed over all the random, person-specific coefficients. The integral has no closed form solution, and in practice simulated maximum likelihood must be used to maximize the unconditional likelihood with respect to the multivariate normal parameters, β and Σ .² Unfortunately, the necessary calculation of multi-dimensional integrals greatly increases the computational burden for these models over the conditional logit model. This can pose a problem for estimating models with many alternatives. As with the joint SP-RH conditional logit, this same model structure can be extended to models that combine data from multiple SP sources.

Note, this parameter structure is the mixed logit analog of Equation 3.8. However, the analog of Equation 3.9, which expresses the SP coefficients as deviations from the RH coefficients, can be made into a very different model. If I replace β by β_{i0} in Equation 3.9, I obtain:

$$\begin{aligned}\beta_{it}^{RH} &= \beta_{i0} + \gamma G_{it} \\ \beta_{is}^{SP} &= \beta_{i0} + \Delta\beta^{SP} + \gamma G_{is} + \Delta\gamma^{SP} G_{is}\end{aligned}\tag{3.19}$$

Compared to Equation 3.7, the SP and RH superscripts are now dropped. The specification in 3.19 represents a dramatically simplified model. It constrains unobserved heterogeneity to be the same across SP and RH scenarios, and assumes that differences in effects across scenarios are captured by the constants $\Delta\beta^{SP}$ and $\Delta\gamma^{SP}$, which are deviations from the RH case that apply in the same measure to all respondents. That is, these deviations only affect the estimates of the mean parameters of the multivariate normal mixing distribution. The variances and covariance terms are now assumed to be equal in SP and RH cases. I make use of this constrained model in the examples that follow.

²In practice, these models are frequently estimated using Halton sequences of quasi-random numbers to calculate multi-dimensional Gaussian integrals. See Train (2000, 2009), Bhat (2001, 2003) and Hole (2007) for additional details.

3.4 Examples: Whites in L.A.FANS

I illustrate models that combine multiple SP observations, as well as SP and RH observations, using the Los Angeles and Neighborhood Survey (L.A.FANS). I use the sample of Whites in L.A.FANS because this sample is large enough to test interesting cases, and because expectations for Whites are more firmly established than expectations for the largest group in L.A.FANS, Latinos. A more detailed discussion of these data is provided in Chapter 2. Individual-level statistics for the samples are shown in Table 2.2. Statistics describing L.A.FANS respondents' RH outcomes are shown in Table 2.4. Descriptive statistics for SP choices are in Table 2.3. I begin with the simpler case of combining multiple SP reports. I focus on the rank SP data from L.A.FANS, illustrating the basic conditional logistic regression approach to test for differences in racial composition effects across rankings. I then illustrate the mixed logit approach for combining these data across rankings and characterizing the degree of preference heterogeneity among L.A.FANS respondents. I move on to demonstrate how to combine data not only across rankings, but from several different SP reports within respondents, using L.A.FANS' "ideal" SP data and rank SP data together. Finally, I illustrate techniques for joint SP-RH models of neighborhood choice.

3.4.1 Models of Rank SP Data

When randomly selected adult respondents ranked five hypothetical neighborhoods in L.A.FANS, they implicitly provided data on four separate choice situations.³ The first question to ask of these ranking data is whether racial composition effects are consistent across the rankings. Is there evidence that Whites are averse to Black neighborhoods to the same degree in choosing the first ranked neighborhood as they are when selecting the second, third, and fourth ranked neighborhoods?

To test for consistency, I estimate models of the form presented in Equation 3.10. Table

³The choice situations are: 1) Ranking of the first neighborhood above all others; 2) ranking of the second neighborhood over three, four, and five 3) ranking of the third neighborhood over four and five; and 4) ranking of the fourth neighborhood over the fifth.

3.1 contains the estimated coefficients. The first column displays estimates for a homogeneous effects model that assumes the effects of racial composition do not differ across ranks. The remaining columns display estimated coefficients from a heterogeneous effects model that permits the effect of racial composition to differ across ranks. Wald tests of equivalence across ranks reject the null hypothesis of no differences.⁴ This suggests that the model allowing for heterogeneous effects provides a better fit to the data than the model that constrains effects to be the same across ranks. However, the BIC statistics for the two models suggest that the model of homogeneous racial composition effects provides a more parsimonious fit.

Predicted probabilities can help to adjudicate between these conflicting statistical tests, especially given the difficulty of directly interpreting quadratic coefficients. The predicted probabilities reveal if the statistically significant differences translate into substantive differences in choices. I plot predicted probabilities for each rank based on the heterogeneous effects model, and predicted probabilities for all ranks using the homogeneous effects model. Predicted preferences for Asian neighbors are shown in Figure 3.1, Black neighbors in Figure 3.2, Latino neighbors in Figure 3.3, and White neighbors in Figure 3.4. Looking across these graphs, there are few substantive differences in the predicted probabilities. All the curves suggest that Whites prefer neighborhoods that are at least majority White, desire limited mixing with Asians and Latinos, and are resistant to neighborhoods with high Black representation. Based on the substantive similarities between all the predicted probability curves, I conclude that the homogeneous effects model provides an adequate summary of preferences.

⁴I use Wald tests rather than likelihood ratio tests because the likelihood ratio tests are invalid when observations are clustered within respondents. The Wald tests, on the other hand, make use of the cluster adjusted covariance matrix.

Accommodating Heterogeneity: The Rank SP Case

The above models assume that preferences are homogeneous among Whites. In that sense, they represent the preferences of a “typical” White resident of Los Angeles. A mixed logit specification, similar to that shown in Equations 3.17 and 3.18, relaxes this homogeneity assumption. I estimate mixed logit models of rank SP vignette responses, assuming consistent preferences across rankings within respondents, but allowing for variation in preferences across respondents. I also present the estimates of the analogous conditional logistic regression for the sake of comparison. Table 3.2 contains the estimated parameters. The first panel shows the estimated means of the coefficients (or simply the coefficients in the conditional logit case). The second panel displays the estimated standard deviations of these coefficients. The first column in Table 3.2 contains the conditional logit results. The second column presents mix logit estimates from a model that assumes that each respondent’s preference coefficients are drawn from independent univariate normal distributions. This allows within respondent preference consistency across rankings, but assumes no association between, for example, a person’s preference for Black neighbors, and his preference for Latino neighbors. The third column presents estimates for a model that relaxes this latter assumption, allowing cross-group association in preferences.

The BIC statistics and the significance of several standard deviation estimates in Table 3.2 all favor models that allow for preference heterogeneity. However, the BIC statistics prefer models that constrain preference variation to independent normal distributions for each racial composition component. Compared to the conditional logit model, the “mean” coefficient estimates from the mixed logit model with univariate normal mixing distributions have larger magnitudes, and these estimates are larger still in the multivariate normal mixed logit case. Also noticeable, the estimated standard deviations in the univariate mixed logit

case are much smaller than in the multivariate normal case.⁵

What of covariation among preferences? Is there evidence that Whites averse to Black neighbors also resist neighborhoods with high Latino representation? Table 3.3 contains estimates of the covariances of racial composition preferences across neighbor groups.⁶ The significant positive covariance between proportion Latino and proportion Black, as well as proportion Latino squared and proportion Black squared, suggests that Whites who are averse to Black neighbors are, indeed, averse to Latino neighbors as well.

The estimates in Tables 3.2 and 3.3 can also be used to derive estimates of preference parameters for individual respondents in the sample (Revelt and Train 2001, Train 2009 Ch. 11). I plot predicted probabilities of neighborhood choice for the mean parameters from the multivariate normal mixed logit model presented in Column 3 of Table 3.2, as well as predicted probabilities for a random selection of 20 Whites from the L.A.FANS sample. Figures 3.5, 3.6, 3.7, and 3.8 depict these predicted probabilities as functions of neighborhood proportion Asian, Black, Latino, and White, respectively. Overall, these predicted probabilities suggest that most Whites have high probabilities of moving into neighborhoods that are at least majority White, but many Whites do prefer neighborhoods where there is at least some intermingling with other racial groups.

3.4.2 Joint Models of Rank SP and Ideal SP Data

Perhaps it is not too surprising to find that Whites in Los Angeles exhibit roughly the same preferences across ranks in the rank SP vignette. However, we might expect to find quite

⁵This points to possible misspecification of the mixing distribution. Additional mixing distributions could be considered, including log normally distributed coefficients. For example, it might be reasonable to constrain the squared terms to be negative log normally distributed. Doing so would force these terms to account only for aversion to neighborhoods with very low or very high representation of some groups. Preferences for greater exposure to a particular group might be better represented by the linear terms alone.

⁶A cursory examination of the diagonal elements of the table reveal standard errors and significance tests that are at odds with those presented in Table 3.2. This is likely because the delta method was employed to estimate standard errors in both Table 3.2 and Table 3.3, but different transformations of the underlying estimated parameters are needed to obtain the standard deviations and the variances. Ideally, a different method would be used to estimate the standard errors. Bootstrapping is one approach, but one that is infeasible for the current case because the estimate of a single mixed logit model is computationally taxing on its own.

different preferences in the ideal SP case, because of the difference in the vignette task and the phrasing of the question. To compare ideal SP and rank SP outcomes, I estimate joint conditional logit and mixed logit models of Ideal SP and RH. Model estimates are presented in Table 3.4. In contrast to the previous models, I include only linear racial composition terms, excluding the quadratic terms, to avoid an explosion in the number of estimated parameters as I test mixed logit models.

I estimate two different specifications for the conditional logit model, corresponding to separate rank SP and ideal SP effects, as in Equation 3.8 (Column 1 of Table 3.4), and a specification treating ideal SP coefficients as deviations from the rank SP coefficients, as in equation 3.9 (Column 3 of Table 3.4). As previously stated, the log likelihoods for these two models are equivalent, but the parameters have different interpretations. The “deviation” specification (Column 3) is helpful because it provides immediately interpretable coefficients. The ideal SP effects in this specification are represented as deviations from the rank SP effects.

The conditional logistic regression coefficients are all negative, suggesting that Whites are averse to neighbors from all racial out-groups. In the rank SP case, Whites are most averse to Black neighbors, followed by Latinos and then Asians, with the latter aversion not statistically significant. The negative ideal SP coefficients in column 3 all point to the fact that Whites express a stronger aversion to non-Whites in the ideal SP vignette. The aversion to Black neighbors remains strongest, but Whites’ show roughly equal aversion to Asians and Latinos in the ideal SP case (judging by the coefficients in Column 1 of Table 3.4).

I also estimate two different mixed logit specifications, one corresponding to Equation 3.7 (Column 2) and one corresponding to Equation 3.19 (Column 4). The mean racial composition effects for the mixed logit models largely conform to the conditional logistic regression results, with Whites proclaiming aversion to all non-white groups, and particularly strong aversion to Blacks. The effect sizes are a touch larger than the conditional logit case,

but the overall consistency between the mixed logit estimates and the conditional logit estimates is encouraging.

The main advantage of the mixed logit specification is that it allows me to assess the degree of preference heterogeneity among Whites in Los Angeles. The estimated population standard deviations in preferences are presented in the second panel of Table 3.4. The model in Column 4, by construction, constrains the standard deviations to be the same in the ideal SP and rank SP cases. The model in Column 2 relaxes this constraint. While the BIC statistics suggest that the constrained model provides a more parsimonious fit, the unconstrained model allows for additional interpretation.

The standard deviations are larger in the rank SP case than they are in the ideal case. This suggests more preference heterogeneity in the rank SP case. This heterogeneity could result from two phenomena. First, Whites might have a narrower range of responses in the ideal SP case, either because of social desirability bias, or because Whites imagine a more racially harmonious and equitable world when asked about their ideal neighborhoods. Second, Whites might have struggled to rank the neighborhoods in the ranking vignette. The ranking task is complicated on its own, and due to the vignette randomization, there may have been only small compositional differences between some of the hypothetical neighborhoods. These factors together could have induced more between individual variability in Whites' responses in the rank SP case relative to the ideal SP case.

Finally, the mixed logit model presented in Column 2 of Table 3.4 also provides estimates of associations between preferences both within and across SP scenarios. Table 3.5 contains the estimated covariance matrix for the mixing distribution. Across rank SP and ideal SP, all of the covariances are positive and significant. This suggests two conclusions. First, aversions to Asian, Latino, and Black neighbors are inter-related. Los Angeles Whites who are averse to Black neighbors also want to avoid Latino and Asian neighbors, and vice versa. This finding holds in both the rank SP and ideal SP vignettes, and adheres to the previous finding from the rank SP data alone. Second, the positive covariances between coefficients

in the rank SP and ideal SP cases signal that responses are at least partly consistent across these scenarios. While these two ways of assessing preferences might be inducing slightly different mean responses, individuals are generally not reversing themselves. Those with higher aversion to Black neighbors in the rank SP data also express higher aversion to Blacks in the ideal SP data, and likewise for Asian and Latino neighbors.

3.4.3 Joint Models of RH and Rank SP

The last example combines residential history and residential preferences data. I use the L.A.FANS rank SP data that I have used across all the previous examples. The RH data are identical to those I analyzed in Chapter 2. I put individuals' RH responses on a discrete time calendar, discretizing the calendar by quarters to pick up as many moves as possible. I note where respondents were living at the end of each interval. I designate the residential location at the end of each interval as the “chosen” alternative, and assume that respondents selected this alternative from a full choice set of all possible Los Angeles County neighborhoods, as given by 2000 United States Census tract boundaries.

I take an approach similar to the one I employed in the previous example. I estimate two parameterizations each of joint SP-RH conditional logit and mixed logit models. One parameterization treats the rank SP data racial composition effects as deviations from the RH effects, and another treats the SP and RH effects separately. As before, the conditional logit specifications are statistically identical, but the mixed logit specifications are not. The differences will be highlighted below.

The RH data demand a richer set of predictors than racial composition, which were modeled as the sole determinant of neighborhood choices in the SP case. In the real housing market immobility, neighborhood proximity, income, education, and housing tenure, among myriad factors, influence neighborhood choices. First, I account for period-to-period immobility, or the tendency for people to “choose” their own houses, by including an alternative in each choice set representing each respondent's own housing unit. I then distinguish this

alternative from the others with a dummy variable. I also include variables to identify the distance of destination tracts from each respondent's origin tract. I account for selection into the migration stream by interacting the own house dummy variable with the respondent's age, marital status, home ownership status, and education. Finally, to deal with sorting along socioeconomic lines, I include interactions between the respondent's income, education, and home ownership status and the levels of income, education, and home ownership in potential destination neighborhoods. I specify these predictors all as fixed, excluding the possibility of effect heterogeneity.

The estimates of racial composition effects from joint SP and RH models are presented in Table 3.6. The first two columns contain estimates from models that treat the SP and RH effects separately, and the last two columns contain estimates from models that treat the SP coefficients as deviations from the RH coefficients. The SP coefficients in the conditional logit models (Column 1 and Column 3 of Table 3.6) conform to the results from the previous example: There are no significant differences in the effects of racial composition between the rank SP and RH cases. In both cases, Whites appear averse to non-White neighbors, with a particular aversion to Black neighbors.

In broad strokes, the mixed logit results (Columns 2 and 4) and conditional logit results (Columns 1 and 3) agree, at least in estimates of the mean parameters. The mixed logit models in both Column 2 and Column 4 of Table 3.6 show that Whites avoid non-White neighbors, especially Blacks, in their actual neighborhood choices. These choices largely agree with choices made in the rank SP vignette. Column 3 and Column 4 present models in which the SP preference coefficients are treated as deviations from the RH coefficients. These deviations are not statistically significant from zero ($p < 0.05$), revealing that Whites' actual neighborhood choices match their stated preferences.

The mixed logit models provide additional nuance concerning the distribution of preferences in the population. Column 2 of Table 3.6 contains estimates of a mixed model that allows for separate, but correlated, SP and RH preference distributions. The first finding of

note: the estimated standard deviations of RH preference parameters in the population are both smaller than the corresponding SP standard deviations, and not significantly different from zero. This suggests that despite statements revealing substantial variation in preferences among Whites, their real world neighborhood choices exhibit little deviation away from a mean tendency to avoid neighborhoods with high non-White representation. Table 3.7, which contains the estimated covariances of preferences within and across scenarios, provides additional insight. Notably, none of the variance and covariance terms that involve the RH racial composition effects are statistically different from zero. This suggests that not only are there no significant covariances within the RH scenarios, but also, there is no association between stated preferences and actual behaviors in the housing market. Whites appear to move to White neighborhoods and avoid non-White neighborhoods regardless of their preferences for inter-racial contact.

This result could emerge from a number of sources. First, the sample of 300 Whites in L.A.FANS is relatively small and these respondents reported few moves. There may be too few longitudinal observations and too few moves in these data to make definitive statements about the association between Whites' preferences and their housing market behaviors. Monte Carlo simulation experiments could be used to investigate the related statistical power issues. Second, there are a number of unobserved factors, potentially correlated with racial composition, that could both affect Whites' neighborhood choices and operate independently of racial composition preferences. Family and friendship ties are one possible factor. If Whites' relatives and friends are also White, and tend to live in neighborhoods that are majority White, and Whites try to locate in neighborhoods close to relatives and friends, this could drive Whites toward White neighborhoods and away from non-White neighborhoods. Finally, Whites might face some perverse biases when making housing choices. Real estate market actors might stereotype all Whites as desiring White neighborhoods, and may subtly steer Whites away from more diverse neighborhoods, even those Whites with tolerant preferences. Whatever the case, the current crop of results suggest that Whites' exposures

to non-White neighbors is over-determined by factors outside of preferences—there is little space for preferences to affect neighborhood choices.

3.5 Conclusion

This chapter has developed and implemented discrete choice methods for comparing stated, race-based neighborhood preferences (SP) to real world residential histories (RH). I outlined the two data types necessary to make these comparisons. I then laid out a conditional logistic regression approach to combining and comparing these two types of data. I also discussed an elaboration of the conditional logit model, the mixed logit. The mixed logit approach is more computationally demanding, but has the benefit of explicitly tying together multiple observations within respondents, while also characterizing the degree of preference heterogeneity in the population. This heterogeneity may be critical to understanding how micro preferences translate into macro segregation outcomes (Xie and Zhou 2012), and, consequently, what the future holds for levels of racial residential segregation in the United States.

The discrete choice methods I describe here have three important advantages over previous, primarily linear regression based approaches to describing the relationship between racial composition preferences and residential mobility. First, discrete choice models explicitly account for the sets of possible neighborhoods to which people move in the RH case, or from which they pick in the SP case. Models that ignore substantial differences in neighborhood “choice sets” across these data types may misstate the relationship between preferences and real-world choices. Second, the discrete choice approach allows an analyst to account for neighborhood preferences and constraints that act on several dimensions at once. This innovation is important in a multi-racial and multi-ethnic world in which neighborhoods are stratified across several racial and ethnic dimensions. In my example, I simultaneously consider the preferences of Los Angeles Whites for Asian, Black, Latino, and White neigh-

bors in a single model. Previous studies examined preferences for and exposures to Asian, Black, Latino, and White neighbors as independent outcomes. Models that can account for multiple neighborhood factors at once are also important for understanding migration in the RH case, where not only do racial and ethnic considerations affect migration, but also price constraints, housing needs, and proximity influence choices of neighborhood destinations. Finally, the discrete choice methods I develop here can be applied directly and without ambiguity to simulation models of segregation. These models continue to be popular in the segregation literature, but with a few exceptions (Bruch and Mare 2006; Xie and Zhou 2012), they are rarely directly grounded in empirical work. In the next chapter I use estimates from discrete choice models presented in Chapter 2 to produce Schelling simulations of segregation processes. In particular, I examine whether the differences between stated preferences and residential history I uncovered in Chapter 2 carry implications for processes and levels of segregation.

Chapter 4

Do Stated Preferences and Residential Histories Imply the Same Levels of Segregation? Testing for Divergence Using a Schelling Model

Douglass Massey and Nancy Denton's *American Apartheid* (1993) made a strong argument that, in the main, housing market discrimination maintains the extreme levels of segregation between Blacks and Whites in the United States. Indeed, members of minority groups face worse treatment than otherwise similar Whites when they seek housing (Turner et al. 2002; Ross and Turner 2005; Yinger 1995). This discrimination appears to extend beyond the point of contact, and into back room mortgage lending practices (Reibel 2000; Williams et al. 2005; Feagin and Sikes 1994).

However, recent racial segregation studies have diverted attention away from practices of housing market discrimination and toward race-based residential preferences. This shift has been motivated by a number of findings and developments. First, the race-based residential preferences of Whites, namely their desire to avoid contact with Blacks, can impel White flight from neighborhoods to which Blacks and other minorities migrate, hollowing out neighborhoods' economic bases and exacerbating segregation and its deleterious effects (Coleman 1975; Frey 1979; Crowder 2000; Crowder and South 2008; Boustan 2010). Second, simulation models of segregation, first proposed and applied by Thomas Schelling (Schelling

1971), have shown how segregation can develop under the influence of preferences alone, absent any institutional discriminating mechanisms (Fossett 2006a,b). These models are particularly compelling for sociologists because they offer a valuable framework for drawing out the complex micro-macro linkages between individuals' and families' behaviors on the one hand, and aggregate, population level patterns of settlement on the other. Finally, new data sources have emerged that provide direct assessments of preferences, most notably in the Detroit Area Study (Farley et al. 1978), and then the Multi-City Study of Urban Inequality (Farley et al. 1993, 1997). The combination of population-based data concerning preferences, a dearth of similar data about acts of discrimination, and compelling theoretical models, has created a headwind for studies that would link segregation to discrimination. The result has been either explicit or implicit acceptance that preferences predominate in driving patterns of macro-segregation.

But is the conclusion that preferences, and not institutional discrimination or perhaps other race-related factors, do most of the work in sustaining segregation justified? Schelling simulation models of segregation, upon which claims about the segregating tendency of preferences are often predicated, make substantial simplifying assumptions about the shape of those preferences (Schelling 1971; Fossett 2006a; Fossett and Waren 2005; Clark and Fossett 2008). These assumptions tend to yield high levels of segregation. But levels of simulated segregation vary substantially across assumptions about functional form, randomness, and heterogeneity (Bruch and Mare 2006; Van De Rijt et al. 2009; Bruch and Mare 2009; Xie and Zhou 2012). Models that adopt more flexible and empirically grounded views of preferences can yield lower levels of segregation than models that rely on simpler preference assumptions. This suggests that while preferences may set a floor for levels of segregation, discrimination or other race related factors can also contribute to extant patterns of segregation.

Even the simulation models grounded in empirical data are frequently based on stated preference (SP) data, not actual housing market experiences. In SP data people respond to hypothetical neighborhoods and identify the neighborhoods in which they would most like

to live. But, as Chapter 2 has shown, what people say they want in a neighborhood does not always align with what they get. Data on residential histories (RH) show that Latinos in Los Angeles, and to a lesser extent Blacks, are partially thwarted in gaining access to the neighborhoods that they prefer, at least in terms of racial composition. This suggests that simulation models should consider two versions of preferences: stated preferences based on what people say they want, and revealed “preferences” based on the choices people make in the real housing market. These two characterizations of how race influences neighborhood choice may imply quite different levels and patterns of segregation.

This chapter asks whether levels of segregation that obtain when people act only on their stated preferences match those that manifest when their race-based residential choices mirror those observed in a real housing market. If the answer is “yes” then those who would attribute observed levels of segregation mainly to preferences may have it right, and institutional discrimination may presently play a minor or negligible role in sustaining segregation. If the answer is “no”, then other race-related factors, potentially including discrimination, may yet be culprits. I investigate this question using modified Schelling models of segregation. In one set of models, I simulate segregation in the case where people migrate based on their stated preferences. I refer to these as stated preference or SP simulations. In another set of models, I examine the amount of segregation that results when migration is guided by the effect of racial composition on neighborhood choices in the real housing market. I refer to those as residential history or RH simulations. SP simulations depict the segregative propensities of “pure” preferences. RH simulations reflect the combined effects of racial preferences, non-racial preferences, and housing market constraints, albeit net of a set of socioeconomic factors that influence residential outcomes: income, education, and housing tenure.

In the remainder of the chapter, I review findings from previous research that has used Schelling models to investigate the link between preferences and segregation. I outline the specification of the Schelling models I employ in my analysis, and discuss the segregation measures I use to evaluate levels of segregation across models. I then present the results from

the Schelling models. I evaluate the differences between segregation indices in SP and RH simulations, and also examine which groups account for differences across sets of simulations. Finally, I conclude and suggest directions for future research.

4.1 Background: Schelling Models and the Importance of Preferences

Schelling models (Schelling 1971) have become a work horse for understanding how individual behaviors interact to generate residential segregation. A Schelling model is an example of an agent-based model (ABM). ABMs posit a set of actors (“agents”) following behavioral rules thought to produce an aggregate, social outcome. The models attempt to “grow” macro-level phenomena from the bottom up, in the process testing hypotheses about what behaviors are sufficient or necessary to generate particular macro-level patterns, and under what system conditions (Axelrod 1997; Epstein 2006; Epstein et al. 1996).

Schelling’s ABM examines whether institutional housing market discrimination is necessary, or whether racial preferences are sufficient to generate high levels of between group residential segregation. The original Schelling model places a group of “Black” agents and a group of “White” agents on a two-dimensional grid of housing units, as depicted in Figure 4.1.¹ The agents move around the city according to preferences for neighborhood racial composition, and nothing more.² The original Schelling model stipulates that the two racial groups have different preferences. In the typical simulation, Black agents are willing to live in neighborhoods that are between 0 and 50% Black, while White agents are willing to live in neighborhoods that are between 0 and 50% White. Beyond these ranges, agents leave their neighborhoods and find new homes in more suitable neighborhoods.

Segregation in the Schelling model evolves in dynamic, complex ways. The Schelling

¹As Schelling points out, imputing races to the groups is just one approach. The groups could also represent other social categories, such as gender, class, age and so on (Schelling 1971).

²Neighborhoods are typically defined relative to cells as in figure 4.2, with neighborhood racial composition calculated over a pre-determined set of surrounding housing units.

model is dynamic because it proceeds in time steps and allows agents to change their locations over the course of the simulation. At each time step, agents evaluate their current neighborhoods and potential destination neighborhoods with respect to their preferences. An agent moves to a new neighborhood if that neighborhood better matches its preferences than the current neighborhood. The models are complex because choices by each agent change the contexts in which all other agents make their choices. As agents update their locations in each time step, neighborhood compositions change. This induces other agents to make moves and so on until either all agents are satisfied (static equilibrium), or until the measures of residential segregation cease to change (dynamic equilibrium).

Schelling's simulation model shows that institutional discrimination is not necessary to generate high degrees of segregation. The internal dynamics of the system generate nearly complete segregation, with Blacks living in almost entirely Black neighborhoods and Whites living in almost entirely White neighborhoods by the end of the simulation. Extensive macro-level segregation "condenses" out of micro-level preferences, even though no single agent prefers entirely own-group neighborhoods to mixed neighborhoods, and no exogenous forces compel Blacks or Whites to live in neighborhoods that do not match their preferences.

Increases in computational power have made it possible to interrogate the behavior of Schelling's model under a number of different initial conditions and behavioral assumptions. Initial conditions, like city dimensions and racial mix, can affect the long-run levels of segregation observed in Schelling models (Fossett and Dietrich 2009; Laurie and Jaggi 2003). Other features of the housing market, like housing quality and housing prices, can articulate with racial composition preferences to generate segregation (Clark and Fossett 2008; Zhang 2004; Bruch 2014).

Most importantly for the present chapter, segregation patterns can play out differently when preferences follow functional forms estimated from empirical data. For example, if preferences are continuous functions of percent out-group, as estimated from observed data for Detroit, then less long-term segregation results, as compared to when preferences follow

a discontinuous threshold function, as in Schelling’s original model (Bruch and Mare 2006). This result is moderated by the degree to which preferences are deterministic (i.e., the frequency with which agents choose neighborhoods that they rank the highest). Continuous preference functions can be *more* conducive to high levels of segregation when choices are made more deterministically (Van De Rijt et al. 2009; Bruch and Mare 2009).³

Even continuous preference functions, estimated from individual level data, may give a misleading impression of the set of preferences operating in the population. Approaches used to estimate preference functions often assume that preferences are homogeneous within racial groups, and that agents’ choices are partly random. Another interpretation of evidence from preference data is that preferences are heterogeneous and deterministic within racial groups (Xie and Zhou 2012; Farley et al. 1993, 1978). When agents in Schelling simulations are assigned heterogeneous preferences and each agent follows a deterministic decision rule, long-run levels of segregation can be lower than when preferences are continuous and homogeneous within groups, with agents following probabilistic decision rules (Xie and Zhou 2012).⁴ Clearly within group variations in preferences, not just between group differences, can influence macro-level segregation patterns.

These studies all suggest that preferences are culpable in generating or preserving some level of segregation between racial groups in the United States. However these studies do not prove that preferences, and preferences alone, are responsible for the levels of segregation that we do observe in the United States today, even disregarding the obvious inertia of segregation regimes established in the middle of the 20th Century. The importance of functional form, randomness, and heterogeneity in mitigating long-run segregation in Schelling models suggests that studies that rely on arbitrary assumptions about the shape and distribution of preferences may misstate the tendency of preferences to generate segregation.

³Benenson et al. (2009) discuss a number of ways to implement deterministic or probabilistic migration rules.

⁴These results are somewhat at odds with Clark and Fossett (2008), who report that long-run segregation levels are the same whether heterogeneous or homogeneous preferences are assumed. However, Clark and Fossett do not ground their assumptions about heterogeneity in empirical data, as Xie and Zhou do. Instead, they rely on an arbitrary normality assumption.

In addition, analyses presented in Chapter 2 suggest that the functional form and strength of stated preferences differ from how race and racial composition affects actual migration. In particular, it appears that non-Whites fall short of achieving their desired proximity to Whites. This implies that the force of race in the actual housing market may drive the system towards greater levels of segregation than would be observed if racial groups were actually able to realize their preferences. I thus hypothesize that Schelling models based on SP data will tend to yield lower levels of segregation than Schelling models based on RH data. I further expect that groups who fall short of matching their preferences will endure less segregation in SP simulations than in RH simulations.

4.2 Methods: Schelling Model Specification

I implement a continuous time Schelling residential mobility model to tease out the implications of SP-RH differences for patterns of residential mobility and racial segregation. The continuous time model is a minor modification of the discrete time Schelling models used in much previous research. Constructing this Schelling model requires two sets of specifications: structural and behavioral. The structural specifications define the city in which agents will move—its shape, size, boundaries, and the composition of its population and housing units. The behavioral specifications define how agents in the city respond to these structural constraints, their neighborhood contexts, and the actions undertaken by other agents.

4.2.1 Structural Specification

The structural parameters play crucial roles in determining the trajectories of racial segregation in Schelling models. These structural parameters enter into the city initialization process, which is diagrammed in Figure 4.3. The structural parameters relate both to the geometry of the artificial city, and the initial make up and distribution of the agents who populate the city.

City Geometry. As in the original Schelling model, I assume a square, two-dimensional

artificial city. The city is carved up into a grid of uniformly sized square cells, as shown in Figure 4.1, with each cell representing a housing unit into which an agent may move, conditional on it being vacant. I assume that the city is 100 cells long by 100 cells wide, for a total of 10,000 housing units. The boundaries of the city are fixed. This means that agents considering moves into or out of neighborhoods at the edges of the city grid evaluate neighborhoods with a restricted number of housing units as compared to agents considering neighborhoods in the middle of the grid.⁵

Population Composition. I specify the racial composition of the agent population to be roughly comparable to that of Los Angeles in the early 2000s, with 40% Latino representation, 30% White representation, 15% Black representation, and 15% Asian representation. I assume that preference heterogeneity is perfectly aligned with group membership—preferences differ between groups, but are homogeneous within groups.

Population Distribution. Finally, I specify the distribution of the population across the city. Like Schelling, I assume that agents are randomly distributed across housing locations on the grid. This initial distribution of agents corresponds to an unsegregated state. I leave 20% of the cells vacant, enabling agents to move to new places on the grid without immediately displacing other agents.⁶

4.2.2 Behavioral Specification

Behavioral specifications in a Schelling model define the calculus agents use in determining when and where they will move during the course of the simulation, subject to the constraints and conditions imposed by the structural specifications. Behavioral specifications include both parameters that are specified at the outset of the simulation, and rules that are embedded into the simulation code. Agent vision is a parameter that determines how many

⁵Alternatively, I could specify periodic boundaries. Periodic boundary conditions treat the “square” grid as a two-dimensional torus that is essentially boundary-less. In previous research, these choices tend not to exert a strong influence on the evolution and equilibria of the simulations.

⁶Past research has shown that segregation outcomes are somewhat sensitive to vacancy rates (Singh et al. 2009). In future research, I test the sensitivity of my conclusions to different assumptions about the vacancy rate.

housing units agents include in their definition of a neighborhood. Preferences are represented by coefficients (parameters) which stipulate how agents translate a neighborhood's racial composition into a neighborhood rating that guides migration decisions. How agents time their migration is determined by rules contained in the simulation code. Rules also determine how neighborhood ratings are translated into moves.

Agent Vision

The agent vision parameter determines how many and which cells agents include in their neighborhood definitions, and thus when determining each neighborhood's racial composition. I assume that agents can see neighbors for two cells in any cardinal or diagonal direction around a candidate cell. This vision corresponds to a 5×5 Moore neighborhood around cells that agents consider as potential residences, as shown in Figure 4.2. I choose this configuration because it roughly corresponds to the size of the hypothetical neighborhoods L.A.FANS responders evaluated during the neighborhood preference vignettes.⁷

Functional Form of Preferences.

Agents' preferences are given by their group memberships. I assume that agents only evaluate neighborhoods on the basis of their own group membership and neighborhood racial composition. In accordance with the discrete choice models developed in Chapter 3, and employed in Chapter 2, each agent, i , assigns a utility to each neighborhood, j , in migration instance t , based on a second degree polynomial in the racial composition of those

⁷In general, it doesn't appear that equilibrium levels of segregation differ much based on reasonable assumptions about the shapes that agents use to define neighborhoods. However, levels of segregation can differ substantially depending on the size of the neighborhoods agents evaluate (Fossett and Dietrich 2009; Singh et al. 2009). In future work, I will test whether the conclusions I present here stand up in the face of different assumptions about agent vision.

neighborhoods:

$$\begin{aligned}
V_{ijt}^{RH} = & \beta_{i1}^{RH} LATINO_{jt} + \beta_{i2}^{RH} LATINO_{jt}^2 + \beta_{i3}^{RH} BLACK_{jt} + \beta_{i4}^{RH} BLACK_{jt}^2 \\
& + \beta_{i5}^{RH} ASIAN_{jt} + \beta_{i6}^{RH} ASIAN_{jt}^2 + \beta_{i7}^{RH} WHITE_{jt}^2
\end{aligned} \tag{4.1}$$

I specify the racial composition preference parameters, β , based on results obtained for L.A.FANS respondents in Chapter 1. These parameters are displayed in Table 2.5. I test three distinct sets of β s to test the implications of SP-RH differences for macro patterns of residential segregation. First, I test β s that reflect the preferences implied by responses to the rank SP vignette. Second, I test β s that correspond to preferences observed in the ideal SP vignette. Third, I test “preferences” based on the racial composition coefficients estimated from L.A.FANS respondents’ residential histories. Over the course of the simulation, I assume that preferences are static within agents.

Timing

I use a continuous time adaptation of the Schelling model to generate migration during simulation. The agent migration process for this continuous time model is diagrammed in Figure 4.4. The continuous time model runs a timer, which I take to be in units of months.⁸ The model begins at month zero, and then proceeds forward in a continuous time until the model is terminated.⁹

The model begins by randomly assigning each agent to a cell (Step 0 in Figure 4.4). Agents “move” to this cell at the beginning of the simulation and become residents (Step

⁸The clock is arbitrary, and its units can be chosen to suit the analyst’s modeling goals. I use months because this seems a reasonable time scale for housing choices. Note, the timer runs as fast as the model can be computed. It is an organizing mechanism for the computations.

⁹To the best of my knowledge, all previous Schelling models have used a discrete time implementation. Discrete time Schelling models proceed in a series of steps. At each step, a single agent is selected as a potential mover, and either remains in place or moves to a new neighborhood, in accordance with the rules of agent behavior and the attributes of the agent’s origin neighborhood and the available neighborhood alternatives. The model is stepped until all agents are satisfied, according to their preferences, or until an equilibrium level of segregation is attained.

1). Upon becoming a neighborhood resident, the simulation assigns a randomly generated clock to each agent. The clock is drawn from an exponential distribution, with a mean six month duration. This clock determines the wait time until the agent's next housing search (Step 2). At every moment during the simulation, each agent is on one of these timers. Because each agent has its own randomly generated timer, the ordering of moves is effectively randomized.¹⁰

Mapping Preferences onto Moves

When an agent's clock/timer expires, the agent immediately transitions into a housing search (Figure 4.4, Step 3). The housing search begins with agents assembling a choice set of 20 randomly selected empty housing units. To this choice set, the agent adds its own housing unit. With this choice set in hand, the agent then proceeds to the neighborhood choice stage (Figure 4.4, Step 4).

The neighborhood choice involves several calculations. First, the agent characterizes each housing unit in the choice set according to the racial composition of the surrounding neighborhood. Using the coefficients shown in Table 2.5 and the functional form in Equation 4.1, the agent assigns a utility to each neighborhood in the choice set (Figure 4.4 Step 4(a)). Finally, I use the conditional logit formula in Equation 4.2 to transform neighborhood utilities into choice probabilities (Figure 4.4 Step 4(b)):

$$P_{ijt} = \frac{\exp(\theta V_{ijt})}{\sum_{k \in C_{it}} \exp(\theta V_{ikt})} \quad (4.2)$$

¹⁰In other versions of the model, I allow other events besides timer expiration to push agents into a housing search. In particular, I allow a change in the neighborhood composition to spur a move into the housing search phase. This leads to essentially the same equilibrium levels of segregation, but the convergence is faster in model time and slower in processor time. Many agents are in sub-optimal neighborhoods at the beginning of the simulation, with some moving quite early on. Each of these early stage moves spurs many other agents to reconsider their residential locations earlier than they would have otherwise. These agents often find that a different neighborhood would be more suitable. However, moves induced in this way diminish as the simulations advances and agents increasingly find themselves in satisfactory neighborhoods.

Using a probabilistic expression to dictate agent choices introduces an extra degree of randomness into the evolution of the system, besides the randomness in the initial distribution of agents and the timing of their moves. The conditional logit probability expression places non-zero choice probabilities on all neighborhoods in the choice set. Even neighborhoods that are highly undesirable according to the utility function have a chance, albeit slim, of being selected during migration. This means that agents occasionally select neighborhoods that provide lower utility than other neighborhoods in their choice sets. I tune the amount of randomness in agents' choices using the parameter θ , which is sometimes referred to as the scale parameter. θ is a positive or zero valued real number. If $\theta = 0$, then each neighborhood in the choice set is given the same choice probability, and neighborhood choices are made entirely at random. At higher values of θ , choices are made more deterministically, with greater choice probabilities assigned to neighborhoods that provide the highest utilities. In the limit, $\theta \rightarrow \infty$, choices are made entirely deterministically, and agents always choose the neighborhood that provides the highest utility. The scale parameter can play an important role in determining equilibrium levels of segregation in Schelling models that rely on probabilistic choices (Van De Rijt et al. 2009; Bruch and Mare 2009). I examine the behavior of both ideal SP, rank SP, and RH based Schelling models across a range of θ values to ensure that my results are not sensitive to my specification of θ .¹¹

I use a multinomial random sampler to assign agents to neighborhoods in their choice sets based on the above expression for the choice probabilities (Figure 4.4 Step 4(b)). Once a neighborhood is selected, the agent then moves to the new cell, becomes a resident, and the cycle begins again (Figure 4.4 Step 1). Steps 3 through 4 in the migration process all happen simultaneously from the perspective of the model. I run the model for 20,000 time steps (i.e., 20,000 months), to ensure that the model attains equilibrium, but in general, approximate equilibrium is reached in approximately 5,000 time steps.

¹¹I also examined a model in which agents make deterministic housing choices, i.e., in which agents always choose to reside in neighborhoods that provide highest utility values. The results obtained from these models are in close agreement with the results obtained when using θ values in the range 5-10. Qualitative snapshots of these deterministic simulations are presented at the bottom of Figures 4.7 and 4.8.

4.2.3 Measuring Segregation

I assess the degree of segregation obtaining in SP and RH based Schelling models using two measures of segregation. Theil’s entropy index provides a single, multi-group summary measure of the degree of segregation that is amenable to decomposition. I also produce more traditional, pairwise dissimilarity indices. These indices identify which groups are more segregated across SP and RH based models. In my analyses, I evaluate the evolution of these indices over the run time of each model, from 0 to 20,000 model time units. I also compare these values at approximate equilibrium, 10,000 model time units into each simulation.

Theil’s entropy index, H , is a valuable measure of multi-group segregation with a number of unique and desirable properties (Theil and Finizza 1971; Theil 1972; Reardon and Firebaugh 2002). It provides a single, summary measure of the evenness with which different groups are distributed across the city. Fundamentally, the entropy index compares the representation of groups at the population level with the representation of those groups in each neighborhood unit. Higher entropy scores can result when many neighborhoods’ racial compositions fall out of line with the population composition, when some neighborhoods greatly deviate from the city’s racial composition, or both.

To calculate the entropy index, first I divide the city up into J exhaustive and mutually exclusive neighborhoods, indexed by j . I assume that the neighborhoods are fixed 5×5 Moore neighborhoods dividing the 100×100 city geography into a grid of 400 neighborhoods. I index the Latino, White, Black, and Asian groups ($M = 4$) in the simulation by m . I then calculate entropy scores for the city as a whole (E), and for each neighborhood (E_j). The city’s entropy score is given by

$$E = \sum_{m=1}^M \pi_m \ln \left(\frac{1}{\pi_m} \right) \quad (4.3)$$

and each neighborhood's entropy score is given by

$$E_j = \sum_{m=1}^M \pi_{jm} \ln \left(\frac{1}{\pi_{jm}} \right) \quad (4.4)$$

where π_m is the representation of group m at the city level and π_{jm} is the representation of group m in neighborhood j . The entropy index is the population weighted mean of the relative deviation of each neighborhood's entropy score from the city's entropy score:

$$H = \sum_{j=1}^J \frac{t_j}{T} \frac{(E - E_j)}{E} \quad (4.5)$$

where t_j is the total population in neighborhood j , and T is the total population of the city. The entropy index varies from 0 to 1, with a value of 0 indicating no segregation, and a value of 1 signaling complete segregation of groups. A nice feature of the entropy index is that it can be easily decomposed into the contributions of each group. Substituting Equations 4.4 and 4.3 into 4.5 and reorganizing yields:

$$H = \sum_{m=1}^M \sum_{j=1}^J \frac{t_j}{TE} \left(\pi_m \ln \left(\frac{1}{\pi_m} \right) - \pi_{jm} \ln \left(\frac{1}{\pi_{jm}} \right) \right) \quad (4.6)$$

I make use of this decomposition to explore which groups contribute the most to segregation differences between SP and RH based simulations.

One drawback of the entropy index is that it does not depict which particular groups are segregated from which other groups. If, based on a decomposition, it appears that Whites contribute the most to the entropy index, is this because they are highly segregated from Blacks, Latinos, Asians, all three, or some combination of the three? To answer these questions, I turn to the index of dissimilarity, a classic measure that has been used extensively

across studies of racial segregation in the United States (Taeuber and Taeuber 1965, 1976; Massey and Denton 1993; Timberlake and Iceland 2007). For two groups, m and n , the dissimilarity index across neighborhoods, indexed by j , is given by:

$$D_{mn} = \frac{1}{2} \sum_{j=1}^J \left| \frac{t_{jm}}{T_m} - \frac{t_{jn}}{T_n} \right| \quad (4.7)$$

T_m and T_n are the total counts of the groups m and n in the city population, respectively. t_{jm} and t_{jn} are the counts of the respective groups in neighborhood j . The dissimilarity index ranges from 0 to 1. A value of zero corresponds to a state of complete integration, whereby the representations of groups m and n in each neighborhood matches that of the city as a whole. A value of 1 indicates complete segregation, where no members of either group live in neighborhoods populated by the other group.¹² The inclusion of four racial groups in the simulations yields six pairwise dissimilarity indices: Latino-White, Latino-Black, Latino-Asian, Black-White, Asian-White, and Black-Asian.

4.3 Results

I simulate the above Schelling model for ten values of the scale parameter and three sets of coefficients—rank SP, ideal SP, and RH. For each combination of the scale parameter and preference coefficients, I perform 10 simulations, each with a different random seed, yielding a total of 300 simulations.

There are several outputs from these models. I first present graphs of segregation statistics as a function of time. I appeal to these graphs to confirm that an equilibrium level of segregation is reached. I pair these graphs with snapshots that provide a qualitative view of

¹²Under random assignment to neighborhoods, the expected value of the dissimilarity index is not 0, but instead depends on the city-level representation of the groups (Winship 1977). Achieving a value of zero requires some concerted effort to spread the population evenly. This does not affect the present analysis because the population composition is held constant across analyses. An adjustment would be necessary if I were to compare simulations in which populations had different compositions.

the evolution of the population distribution over the course of simulations. I then focus on differences between SP and RH based simulation. I examine whether equilibrium levels of segregation differ for simulations based on the RH racial composition coefficients and those based on SP coefficients. I also examine whether the results are consistent across different assumptions about the scale parameter, θ . Finally, I examine which groups contribute the most to segregation indices, and to which groups I can attribute SP-RH differences.

4.3.1 Attaining Equilibrium

Plots of both the entropy index and dissimilarity indices over time suggest that the Schelling models reach convergence. Figure 4.5 depicts the convergence of the entropy index. Figure 4.6 depicts the convergence of the dissimilarity indices. With some minor variations, for each set of ideal SP, rank SP, and RH based simulations the Schelling model appears to reach convergence after approximately 5,000 time steps. The models reach segregated equilibria across values of the scale parameter, suggesting that the equilibrium is dynamic—it holds even when agents’ moves are highly random and they remain continuously mobile.

The plots of the entropy index over time also suggest that segregation in RH based simulations exceeds that observed in rank SP based simulations. The same appears to hold true for the ideal SP based simulations, although with the qualification that the scale parameter influences this result. The discrepancy between rank SP and RH simulations on this multi-group measure of segregation does not, however, appear to extend to pairwise measures. Plots of dissimilarity indices, which indicate the degree of segregation between pairs of groups, suggest that levels of Latino-Black segregation are more or less consistent across simulations based on RH and rank SP coefficients. Again, the comparison of dissimilarity indices for ideal SP and RH based simulations hinges on the scale parameter setting.

4.3.2 Qualitative View of Segregation

Visual inspection of the segregated landscape after 10,000 steps confirms qualitative differences between SP and RH based simulations. Figure 4.7 depicts the population distribution for sample RH based simulations. Figure 4.8 depicts the evolution of the population distribution for sample rank SP simulations. Both sets of figures display sample simulations started from the same random seed. This means that each of the simulations pictured begins with the same initial, pseudo-random distribution of agents. The rows of sub-figures correspond to different values of the scale parameter. Columns correspond to time points.

In the RH based simulations (Figure 4.7), large blocks of Whites and Latinos form, even when the scale parameter is held at a relatively low value of 1. At higher values of the scale parameter, large blocks of Latinos also develop. Blacks are largely excluded from Asian and White areas in all cases. At low values of the scale parameter, Blacks are intermingled with Latinos, but at high values of the scale parameter, Blacks are also excluded from Latinos' neighborhoods. However, rather than forming large blocks of uniformly Black neighborhoods, as Latinos, Whites, and Asians do, Blacks tend to occupy thin neighborhood bands in the interstices between Latino, White, and Asian neighborhood blocks.

In the rank SP case (Figure 4.8) segregation barely manifests when the scale parameter is held at the low value of one. At higher values of the scale parameter, noticeable segregation condenses out of the initial random distribution of agents. The patterns of segregation differ qualitatively from those observed in the RH case. First, White and Latino clusters are directly adjacent to each other. In the RH based simulations, White and Latino clusters are separated by bands of Black neighborhoods. Second, the White neighborhoods and Asian neighborhoods are less monolithic. While larger clusters of Whites do appear, smaller pockets of Whites also dot the landscape. In the case of Asians, their behavior bears greater resemblance to that of Blacks in the RH case. There are no large blocks of Asian neighborhoods. Instead, ribbons of Asian neighborhoods compete with strands of Black neighborhoods in wrapping around Latino and White neighborhoods. The slightly more

interspersed patterns in the SP simulations compared to the RH simulations suggest that levels of segregation measured by segregation indices will be lower as well.

4.3.3 SP-RH Differences: Entropy

Next I quantify the degree of difference between segregation in the SP and RH simulations and more carefully assess whether those differences persist across different assumptions about the degree to which choices are made randomly. I present mean entropy and dissimilarity indices for simulations based on different SP and RH racial composition coefficients, and for different values of the scale parameter. Table 4.2 contains the entropy indices and Table 4.1 contains the dissimilarity indices. Because I executed ten randomly seeded simulations for each set of coefficients and each value of the scale parameter, these tables also contain standard deviations. However, relative to the value of the indices, the standard deviations are quite small, and in general t-tests reveal most between scenario differences to be statistically significant.¹³

Table 4.2 confirms that the entropy index is significantly lower in the rank SP case than in either the ideal SP or the RH cases. Segregation in the rank SP case is very low ($H = 0.092$) when the scale parameter is set to 1. In contrast, entropy in the RH case falls significantly above zero ($H = 0.498$) and entropy in the ideal SP case is even higher ($H = 0.739$). The difference between the rank SP and RH simulations declines as the scale parameter is increased, but their relative rankings remain the same. Even at high values of the scale parameter ($\theta = 5$) the rank SP entropy index is lower than in the RH and ideal SP simulations. While the entropy difference when θ is set to 5 is not as dramatic as when θ is set to 1, the segregation in the RH simulations is over 25% higher $((0.839 - 0.660)/0.660 = 0.271)$ than in the rank SP simulations. This gap is not negligible. While this does suggest that segregation is largely due to preferences, this also leaves substantial space for other race-

¹³These standard deviations neglect to treat the uncertainty in the parameter estimates. Future versions of this work will take account of uncertainty in parameter estimates by using the estimated coefficient variance-covariance matrix in conjunction with asymptotic normality assumptions to draw multiple plausible values for the preference coefficients, running simulations for each draw of the coefficients.

related factors to exacerbate segregation.

I confirm that the SP-RH differences are not unique to the scale parameters presented in Table 4.2 by plotting the equilibrium entropy index in the SP and RH cases as functions of the scale parameter. Figure 4.9 confirms that segregation in RH simulations exceeds that in rank SP simulations across a broad range of assumptions about the scale parameter. However, there is crossover in levels of segregation for the ideal SP and RH simulations when $\theta = 2$, after which segregation in the RH simulations exceeds that observed in the ideal SP simulations. In fact, the level of segregation to which the ideal SP simulations converge does not vary substantially across different assumptions about the scale parameter. This suggests that individuals' responses to the ideal SP vignette in L.A.FANS are more certain than their responses to the rank SP vignette. This is not to say that the ideal SP vignette provides a better characterization of preferences, just that the preferences are sharp enough to leave their implications in simulation mostly unaffected by the scale parameter.

In sum, simulations based on stated racial preferences for residents of Los Angeles lead to a more even distribution of groups, and hence lower levels of segregation, than in simulations based on “revealed” preferences derived from residential history data. The implication is that while stated preferences are conducive to some degree of segregation, as previous studies have suggested, they do not, necessarily, square with observed levels of segregation. The inability of some groups to achieve their preferences leads to greater segregation than there might otherwise be, at least when we abstract away from other housing market factors, such as income, wealth, education, and household needs.

4.3.4 Decomposing the Entropy Index

The preceding analysis tells us very little about which groups are more or less segregated in RH as opposed to SP cases. Is the degree of segregation uniformly lessened for Latinos, Whites, Blacks, and Asians, or are there differences across these groups? To answer this question, I decompose the entropy index from SP and RH simulations into the separate

contributions of the four racial groups. I then compare contributions across SP and RH simulations to determine which groups are most responsible for differences in the equilibrium entropy index. The entropy index decomposition is based on Equation 4.6. Table 4.2 contains the result of the decomposition.

Within scenarios, Latino and White groups have the highest contributions to the entropy index. However, differences in the segregation of Whites and Asians contribute most to lower segregation in the rank SP case relative to the RH case. Differences in the distribution of Whites account for approximately 35% of the difference between rank SP and RH simulations ($(0.238 - 0.174)/(0.839 - 0.660) = 0.352$). Differences for Asians account for another 36% ($(0.204 - 0.141)/(0.839 - 0.660) = 0.358$). Latinos, on the other hand, account for approximately 24% of the difference, and Blacks account for only 6%. This result aligns with the previously discussed qualitative results in Figures 4.7 and 4.8. White neighborhood clusters in the SP case are adjacent to a greater variety of neighborhoods, including Asian and Latino neighborhoods. This differs from the RH case, in which White neighborhoods are surrounded by thin rings of Blacks. At the same time, White neighborhood blocks are smaller and more dispersed in the SP case. These all suggest lower levels of segregation for Whites. Meanwhile, condensed Asian neighborhood clusters never form in the SP case as they do in the RH case. Instead, Asians attain greater contact with Latino and White neighborhood clusters, while also intermingling with Blacks. Both sets of figures exhibit large blocks of Latino neighborhoods, but these neighborhoods are adjacent, to various degrees, to both White, Asian, and Black neighborhood clusters, rather than being separated by bands of Black neighborhood clusters. Blacks continue to be arrayed in thin bands around other ethnic neighborhood clusters, but these bands are less continuous than in the RH case.

This does not necessarily imply that it is the behaviors of Asians and Whites that lead to the SP-RH differences. The complexity of social interactions in Schelling models generally precludes such facile conclusions. Naturally, the lower segregation of Whites in the rank SP case, for example, could be due to Whites expressing more tolerant preferences than they

appear to follow in the real housing market, but SP-RH differences in the preferences of Blacks, Latinos, and Asians may also factor into it. Their inability to translate preferences into migration in the RH case may be the source of lesser White segregation. And of course, the lower levels of segregation could be due to the complex interaction between all groups' preferences. Examination of segregation measures for pairs of groups can help to shine additional light on these questions.

4.3.5 Dissimilarity Indices

The entropy index does not resolve which groups are segregated from each other. If Whites are less segregated in the rank SP case than in the RH case, this could be due to greater exposure to Latinos, Asians, Blacks, or some combination of the three. Here, the dissimilarity index offers some purchase. I calculated the dissimilarity index at the 10,000th time step for each set of preferences and each value of the scale factor. Dissimilarity indices for a selection of scale parameter values are shown in Table 4.1. Figure 4.10 plots the dissimilarity as a function of the scale parameter for the six possible racial pairs. I focus on the results for Whites and Asians, because they account for a majority of the difference between the entropy index in rank SP and RH simulations, but the other groups inevitably factor into the conversation. I set aside discussion of the ideal SP data because of concerns about how well these data actually represent the preferences that people bring to the housing market.

The lower segregation of Whites in the rank SP simulations is primarily driven by their greater integration with Latinos and Asians. Of the three dissimilarity indices that involve Whites in Table 4.1—White-Latino, White-Black, and White-Asian—White-Latino and White-Asian segregation indices are substantially lower in the rank SP case. Black-White segregation is only lower in SP simulations when the scale parameter is held at a value below seven. At higher values of the scale parameter, Black-White segregation does not differ between SP and RH simulations. Even when the Black-White dissimilarity is lower in the rank SP case, the gap between it and the dissimilarity index in the RH case is rela-

tively small. This suggests that greater integration of Blacks with Whites is not responsible for lower White segregation in the rank SP case.

Re-examination of the plots of neighborhood choice probabilities discussed in Chapter 2 provides some suggestive evidence that both Asians' preferences and Whites' preferences are leading to their lower segregation in the Rank SP simulations for Whites. Whites state a preference for some mixing with Asian neighbors, assigning higher choice probabilities to neighborhoods with nominal ($\tilde{20\%}$) Asian representation (see Figure 2.10). Their actual neighborhood choices, however, betray a weak tendency to move to neighborhoods with no Asian representation at all. Meanwhile, Whites' stated preferences for White neighbors are weaker than is implied by Whites' actual residential mobility (Figure 2.9). Asians, on the other hand, state both a stronger preference for Asian neighbors, and a slightly stronger preference for Whites as neighbors in the RH data relative to the rank SP data.

For Latino-White segregation, lower levels of segregation in rank SP simulations might be spurred by the greater desire on the part of Latinos to have White neighbors. The predicted probabilities shown in Figure 2.5 show that Latinos have a stronger preference for White neighbors, or at least a greater willingness to live with Whites, in the rank SP data than they realize in their residential histories. However, Latinos' residential histories suggest that they are unable to live with as many Whites as they would like, even controlling for several other factors that might impede Latinos in matching their preferences. Based on qualitative snapshots in Figure 4.8, this greater willingness to live near Whites appears to translate not into mixed Latino-White areas, but rather clusters of White and Latino neighborhood that directly border each other. This stands in contrast to the bands of Black neighborhoods that separate Latino and White neighborhoods in the RH simulations. Shared borders are the source of more Latino-White integration in the SP simulations.

In addition to lower Asian-White segregation, Asian-Black segregation appears to contribute to lower Asian segregation in the rank SP simulations. According to Table 4.1, Asian-Black segregation is lower in the rank SP scenarios across all values of the scale pa-

parameter. It is difficult to attribute this directly to Asians' and Blacks' preferences for Asian and Black neighbors. However, the visual patterns of segregation shown in Figures 4.7 and 4.8 suggest one possible explanation. Absent any change in preferences for Asian and Black neighbors, a stronger tendency for Asians to live near Whites and a weaker tendency to live near other Asians and avoid Blacks might bring Asians into greater contact with Blacks in the rank SP case. This result may also depend on Black desires to live with some Whites, and willingness to live in neighborhoods with some Asian representation.

4.4 Discussion and Conclusion

This paper attempts to determine whether stated racial composition preferences, differing among broadly construed racial groups, conduct populations toward more or less segregation than implied by the effects of racial composition on residential mobility in the real housing market. If people were able to move to the neighborhoods they claim to prefer, would that tend to lead to more or less segregation than we actually observe? To answer this question, I implement Schelling-like agent based models of inter-neighborhood migration. These models stipulate four distinct sets of actors, representing four broad racial groups in Los Angeles County: Latinos, Whites, Blacks, and Asians. Members of these groups populate an artificial city, represented by a two dimensional square grid, and then move about this city according to a set of racial preferences. Rather than assuming a threshold function for preferences, as many Schelling models do, I follow Bruch and Mare (2006) and bestow preferences on groups of agents according to empirical models. In one set of simulations, I assign racial composition preferences to agents based on empirical models of stated preferences (SP) presented in Chapter 2. Notionally, the SP reports represent the racial composition of the neighborhoods where people most wish to live, free from housing market constraints. In another set of simulations, agents draw their racial preferences from empirical models of residential mobility contained in Los Angeles residents' residential histories (RH). These

“preferences” actually represent the interactions between racial preferences and other housing market constraints and preferences, including racial discrimination and the desire to live near family and friends.

I characterize the evolution of these simulations using two sets of segregation measures, one multi-group measure, the entropy index, and dissimilarity indices, which track the level of segregation between pairs of groups. The entropy index shows that, under a broad array of assumptions about the degree of randomness in agents’ housing choices, simulations based on the stated preferences tend to yield lower levels of segregation than simulations based upon residential history data. Conservatively, RH based simulations yield entropy indices that are more than 25% higher than those observed in SP simulations based on the rank SP vignette discussed in Chapter 2. That is, the influence of race in the real housing market tends to generate higher levels of segregation than is implied by preferences alone. Decomposition of the entropy index and examination of dissimilarity indices reveals that lower levels of segregation among Whites and Asians, and to a lesser degree, Latinos, explains much of the SP-RH difference. In particular, levels of Latino-White, Asian-White, and Black-Asian segregation are notably lower in SP based simulations than in RH simulations, especially in simulations based on rank SP data.

The analysis suggests that racial preferences are not the be all and end all of racially segregating factors at work in the housing market. If racial preferences operate alone, they tend to yield segregation, yes, but less segregation than manifests when agents move in accordance with the real housing market effects of race and racial composition. Thus, observed levels of segregation are not fully consistent with, nor are they likely wholly perpetuated by, racial composition preferences alone. Instead, these results agree with Fossett’s (2006a) conclusion: “segregation at present, and likely in the recent past as well, has been sustained by a combination of two sufficient causes—discrimination dynamics and social distance and preference dynamics.” (p. 258) However, Fossett goes on to claim “If this is the case, there is little reason to expect segregation to decline immediately and dramatically following declines

in discrimination because another sufficient cause (namely, ethnic preferences) may still be operating to sustain segregation at high levels.” To the extent that findings presented here are driven by discrimination, they give cause for slightly more optimism. While segregation may well persist in the absence of institutional discrimination, it appears that the abolition of racially discriminatory practices could lead to reductions in segregation of up to 20%, at least in cities featuring an ethnic mix similar to Los Angeles.

To be sure, there are elements left out of the models presented here. First, the simulation models focus purely on race and racial composition, to the detriment of other factors that could contribute to or mitigate segregation. While the RH based racial composition “preferences” I employ in the simulations net out individual and neighborhood level factors like home ownership, education, and income, the simulations themselves ignore the possibility that these factors can interact with race and racial composition in complex ways to produce segregation (Bruch 2014). And, of course, the RH racial composition effects upon which I base my Schelling simulations may not appropriately control for socioeconomic factors, like housing costs or proximity to employment, or social network factors, like the spatial distribution of friends and family, that might explain some of the discrepancies between the racial preferences people express and the effects racial composition has on real housing market decisions.

There are also two methodological issues that deserve further consideration. First, the simulations considered here, and the estimated differences between SP and RH, are based on the mean racial composition parameter estimates derived from conditional logit models of preferences and neighborhood choice. I have ignored the uncertainty in the parameter estimates, but this uncertainty could influence conclusions about the *statistical significance* of SP-RH segregation differences. Second, the simulations also ignore the possibility of preference heterogeneity among respondents nested within groups. However, heterogeneity may play an important role in mitigating macro patterns of segregation (Xie and Zhou 2012). Future work should more carefully consider both uncertainty about parameter estimates,

and the possibility of heterogeneity in preferences among agents.

These qualifications aside, this study represents a new stride in the development of models that bridge the gap between empirical knowledge of housing preferences and racial segregation, on one hand, and simulation-based theoretical understanding of processes of segregation. This study builds on the results presented in Bruch and Mare (2006), and reveals how empirically measured racial composition preferences can generate different patterns of segregation than are observed in highly abstracted simulation models that take relatively un-nuanced views about the shape of preferences. This may provide some encouragement to those scholars who operate with an implicit understanding that what people want and what they get in the housing market are unlikely to match, especially in the case of minority groups. Rather than ceding ground to simulation methodologists, scholars with empirical interests in interactions between preferences and discrimination should think carefully about how they can articulate those concerns with simulation models, and so draw out the implications of processes of housing market discrimination for macro-level segregation outcomes.

Chapter 5

Conclusion

Residential segregation between racial groups has remained a knotty issue in the United States. While recent decades have seen declines in segregation between Blacks and Whites (Logan et al. 2004; Iceland and Sharp 2013), the observed levels of segregation remain quite high. And Blacks continue to endure greater segregation than broadly construed Asian and Latino groups, even as levels of Asian and Latino segregation have remained stable or even increased slightly over the last thirty years. The increase or stability of Asian and Hispanic segregation has also raised the specter of segmented and downward assimilation for Asian and Hispanic groups enduring exposure to homogeneous neighborhoods (Zhou et al. 2008).

Race-based residential preferences are frequently tagged as the most likely suspects in perpetuating racial residential segregation (Fossett 2006a, 2011; Macy and Van De Rijt 2006). This attention has been partly spurred by a recognition that patterns of White flight, presumably driven by Whites' desires to avoid minority residential contact, contributed significantly to the creation of segregated regimes in the middle of the century (Boustan 2010). Increasingly, the popularity of simulation models of residential segregation, which show that observed between group differences in race-based residential preferences alone are sufficient to induce high levels of segregation, have added theoretically incriminating evidence to the indictment of residential preferences. The persistence of notable, albeit varying degrees of in-group preferences among Whites, Asians, Blacks, and Latinos, presented in the preceding chapters and elsewhere (Farley et al. 1997; Charles 2000, 2007) have led some to

express skepticism that reductions in racially biased housing market practices, or reductions in between-group socioeconomic inequalities, will have much effect on levels of residential segregation (Clark 1991; Fossett 2006a, 2011).

A number of scholars have resisted the notion that preferences alone can sustain the high degrees of Black-White residential segregation, not to mention segregation between other groups. Objectors to preference based explanations of segregation have generally fallen into two camps. One camp explicitly advocates for the capacity of other mechanisms, especially housing market discrimination, to sustain segregation (Galster and Keeney 1988; Galster 1987, 1988; Massey and Denton 1993; Massey 2005). The other camp remains agnostic about the causes of observed levels of segregation, but notes that the tendency for preference to generate high levels of segregation may be over-stated (Bruch and Mare 2006; Xie and Zhou 2012).

Scholars who advocate for other mechanisms, besides preferences, have two social facts on their side. First, racial disparities in socioeconomic attainments are still an important feature of the United States and its housing markets. There are persistent gaps between Blacks and Whites in terms of income, education, and wealth. At the same time, even those who reach socioeconomic parity with the White majority face poor treatment in the housing market. Housing market discrimination continues to disproportionately afflict Black and Latino housing seekers (Yinger 1995; Turner et al. 2002; Ross and Turner 2005; Turner et al. 2013). And even those given fair treatment by real estate agents and landlords may find it difficult to gain the financial backing they need to enter White neighborhoods (Ondrich et al. 1999; Reibel 2000). This intellectual camp argues, with some theoretical support (Fossett 2011), that discriminatory housing market regimes, in combination with socioeconomic inequalities, could contribute substantially to observed levels of segregation.

Those who remain agnostic about explanations of segregation highlight the empirical deficiencies in current models that link segregation to preferences. Importantly, the segregative tendency of preferences imputed to “agents” in simulation models is sensitive to the speci-

fication of those preferences. Simulation models directly based on empirical assessments of preferences appear to yield lower levels of segregation than simulation models that adopt caricatured views of preferences, frequently based on threshold functions, that are only loosely tethered to empirical observations (Van De Rijt et al. 2009; Bruch and Mare 2006; Xie and Zhou 2012). These results render it less clear whether extant levels of racial segregation can be attributed to preferences alone. There may yet be a gap between the levels of segregation observed in the real world, and those implied by direct assessments of preferences, and that gap could be occupied by other mechanisms.

Significantly, previous theoretical simulations, even those that give careful consideration to preference functions, have not bothered to examine whether the influence of race and racial composition in real housing markets mirrors people’s stated, race-based residential preferences. That is, simulation studies frequently proceed from the unverified assumption that people get what they say they want when it comes to neighborhood racial composition.¹ Blacks who prefer mixed neighborhoods are assumed to migrate in accordance with those preferences, and Whites who proclaim aversion to non-Whites are likewise able to locate in neighborhoods consistent with those preferences. But this glosses over the arguments made by other scholars that not all groups are able to match neighborhood attainments to preferences. One of the fundamental tenets in the literature on housing market discrimination is that disadvantaged minorities are unable to gain access to neighborhoods that match their preferences.

This dissertation has re-examined the question of whether, and how much, segregation can be attributed to preferences, as opposed to other, race-related forces that operate in the housing market. It has done so in three ways. First, it has developed methods for comparing neighborhood preferences to neighborhood attainments. Second, it has used those methods to analyze empirical data describing preferences and inter-neighborhood migration in Los Angeles. Third, it has simulated segregation processes based on the empirical analyses of

¹A recent exception is Bruch (2014), who uses data on migration to infer “revealed” preferences, and then applies these inferred preferences in simulation models.

preferences and residential histories.

Chapter 2 compared stated preferences for racial composition to the effect of racial composition on actual housing market outcomes (“revealed preferences”). Theories of residential segregation suggest that preferences and neighborhood attainments will be at odds for some groups. This chapter provided tests of these conjectures. Stated preference data, from neighborhood vignettes, and revealed preference data, in the form of residential histories, were obtained from the Los Angeles Family and Neighborhood Survey. I used discrete choice modeling techniques to test whether preferences and actual neighborhood choices matched, and whether some groups matched preferences to actual choices more successfully than others. I accounted for the matching of people to neighborhoods according to income, education, and home ownership using statistical controls at both the individual and neighborhood levels. I also accounted for the spatial contingency of residential mobility by introducing controls for distance migrated.

Results from Chapter 2 showed that Blacks, Latinos, and Whites in Los Angeles are all, to some degree, unable to migrate in accordance their neighborhood racial composition preferences. However Latinos and Blacks endure greater mismatch between their stated preferences and their residential experiences than Whites do. After controlling for sorting on socioeconomic and other non-racial variables, Latinos face the greatest disadvantages. Relative to their preferences, Latinos are over-exposed to other Latinos, and under-exposed to Whites. Los Angeles Blacks are also disadvantaged relative to Whites, but not as much as previous literature on processes of segregation would suggest. Like Latinos, Los Angeles Blacks endure some over-exposure to Latinos, and under-exposure to Whites, but these over- and under-exposures appear to be partly mitigated by introducing controls for nonracial factors.

Chapter 3 developed the methods used in Chapter 2. These methods are extensions of previous methods used to combine data on stated preferences and “revealed” preferences in other contexts, especially in studies of transit choice (Ben-Akiva et al. 1994; Morikawa

1994; Hensher et al. 1998; Swait et al. 1994). Unlike these previous approaches, which tend to assume that stated preference data, based on controlled, vignette experiments, and revealed preference data, based on observed residential histories and the implicit choices contained therein, should match, methods developed in Chapter 3 treat the matching of preferences to behaviors as an empirical question. Problemetizing the preference-migration match is necessitated in the residential mobility literature because of a number of theoretical perspectives that predict a mismatch. Chapter 3 also presents additional techniques, not employed in Chapter 2, to exploit panel data to quantify the heterogeneity of race-based neighborhood preferences within broad racial groups.

Finally, Chapter 4 used the empirical results from Chapter 2 to explore the implications of preference-migration discrepancies for segregation. In Chapter 2, stated preference assessments notionally represented “pure” preferences for neighborhood racial composition. Racial composition preferences “revealed” in Los Angeles residents’ residential histories resulted from the intersection of preferences with housing market constraints, including racial discrimination and social networks, but excluding spatial and socioeconomic constraints for which the analysis employed statistical controls. Chapter 4 presented two sets of Schelling simulations, one set based on residential history data (RH simulations) and another set based on stated preference data (SP simulations). Patterns of segregation obtaining in these two different sets of simulations were then compared.

Results for Chapter 4 showed that SP simulations produced lower levels of segregation than RH simulations. A single summary measure of multi-group segregation, the entropy index, was approximately 25% higher in RH simulations than in SP simulations. Whites and Latinos made the largest contributions to this difference, with Whites and Latinos experiencing higher rates of residential contact with each other in the SP simulations than in the RH simulations. Whites were also more integrated with Asians in the SP simulations. These results implied that, while preferences are conducive to segregation, other race related mechanisms at work in the housing market, perhaps including discrimination and kinship

and friendship networks, tend to increase levels of segregation. Preferences may set a floor for levels of segregation, but this floor may not be so high as to rule out other sources of segregation.

There is much work left to do in understanding how preferences, relative to other factors, influence racial residential segregation. Reflecting the structure of this dissertation, new areas of inquiry can roughly be split into empirical, methodological, and theoretical simulation domains.

On the empirical front, it remains difficult to quantify the relative influences of preferences, socioeconomic inequalities, pre-existing spatial patterns of segregation, and kinship networks, among many candidates, on the persistence of segregation. The main barrier is a general dearth of data that integrates measures of the multiple potential contributors to patterns of segregation, or a complete lack of data on key factors. To begin, with the exception of L.A.FANS, the co-existence of stated preference and residential history observation on the same individuals is quite rare. While several studies, including the Multi-City Study of Urban Inequality, have surveyed preferences across contexts and racial groups, these studies lack a longitudinal residential history component. But the residential history portion is key to understanding how preferences relate to behaviors. Lacking these data in other parts of the United States, it is difficult to see if the results obtained in Los Angeles, and presented in this dissertation, apply to other metropolitan areas in the United States. Are the same patterns of SP-RH discrepancies in evidence? In particular, do Blacks experience larger SP-RH discrepancies in cities with larger Black populations, and longer histories of racial segregation between Blacks and Whites?

The lack of data on acts of discrimination will likely continue to present an obstacle for theories that would tie segregation to housing market discrimination. It is not clear how this issue will be remedied with data. The data on acts of housing market discrimination are likely fundamentally unobservable, and suitable proxies may be exceedingly difficult to develop. Instead, it appears that discrimination will continue to be treated as a residual

explanation: once all other factors that are thought to influence migration are controlled for, the remaining effect of race and racial composition on migration, or the remaining difference between stated preferences and actual residential attainments, might be tentatively attributed to discrimination.

However, the list of other factors, besides discrimination, preferences, and socioeconomic attainments, that might contribute to segregation has not yet been exhausted. I made the argument in Chapter 1 that racially patterned friendship and kinship networks could perpetuate segregation by creating ties of social obligations between people and the frequently segregated neighborhoods in which they are born and raised. The Panel Study of Income Dynamics, however limited its sample, presents one opportunity to trace out the effects of proximity to kin on patterns of inter-neighborhood migration, and thus segregation. The opportunities to inspect the role that kin play in migration are even richer in countries with extensive administrative data sets. For example, in Sweden, researchers now have access to population registry data describing the schooling, residential, and workplace histories of nearly every person who lived in Sweden over the course of two decades, beginning in 1990 (Statistics Sweden 2014). These records also contain identifiers linking individuals to their parents, which can be used to establish the spatial distribution of family members. These variables can then be used in analyses that determine whether neighborhood racial or ethnic composition influences migration net of proximity to kin.

The Swedish registry data are not limited to tracing out the influence of kinship networks on migration. Using these data, it is also possible to tie individuals to schools and workplaces. Considering that proximity to schools and work places may serve as important guides of choices of residential location, this presents an additional opportunity to consider joint models of school choice and residential choice, or of residential choice and workplace choice (Waddell et al. 2007). In the presence of spatial mismatch, i.e., difficult to surmount spatial distances between affordable neighborhoods and places of work, this may be a race-correlated factor that influences patterns of residential segregation by race and ethnicity.

Schools and work places are also potential sources of friendship and acquaintanceship, which may influence residential location decisions. These social networks may impact migration in two main ways. First, people may seek to live with or near friends, thus influencing choices of residential destinations. People may also use these social networks to glean information about residential opportunities. Thus people may move near friends, or friends of friends, because these social contacts provide privileged access to knowledge about residential vacancies. To the degree that schools and workplaces, and by extension friendship and acquaintanceship networks, are sorted along racial and ethnic lines, these networks may explain some portion of the apparent effect of racial and ethnic composition on processes of inter-neighborhood migration, and so account for some of the segregative power of residential choices. Registry data present an opportunity to characterize these networks because they allow individuals to be linked by dint of a shared employer or schooling experience.

On the methodological front, techniques for estimating models that explicitly account for heterogeneous effects and panel data are an emerging frontier. These techniques take a number of different forms, including the latent class logit (Chintagunta et al. 1991; Greene and Hensher 2003; Train 2008) and the parametric mixed logit (Revelt and Train 1998; McFadden and Train 2000; Train 2009). These methods have been applied in the consumer and transit choice literatures, and increasingly there are programs available to estimate these models using popular statistical software packages (Hole 2007; Chang and Lusk 2011; Hole 2011). However, these models have not been applied much in the residential choice literature. The key challenges for these approaches are computational. The models are made more computationally intensive by the necessity of employing EM or simulated maximum likelihood algorithms. Mixed logit models are also unable to accommodate methods of sampling from the neighborhood choice set, an approach that makes traditional conditional logit models much more computationally tractable without inducing bias.

Mixed logit and latent class logit methods deserve attention because they allow for potentially unbiased empirical investigation of a number of theoretically interesting behaviors.

First, they permit the characterization of preference heterogeneity. Heterogeneity on its own can make a large difference in segregation outcomes (Xie and Zhou 2012). Second, these methods can be used to investigate state dependence in residential choices. State dependence occurs when an individual's past choices influence future choices (Heckman 1981; Morikawa 1994). State dependence in residential choices is one potential pathway in the social reproduction of race, as well as of residential segregation. If people partly learn their race-based preferences from prior residential exposures to other groups, then state dependence may be a crucial pathway for continuing between-group disparities in race-based residential preferences, and thus residential segregation. Finally, mixed models also offer means to jointly model neighborhood choices with other spatially situated outcomes, including school attendance, employment and transportation choice (Waddell et al. 2007; Pinjari et al. 2011; Lee and Waddell 2010). Racial stratification along these other concrete, but often overlooked, socioeconomic dimensions may make important contributions to segregation in some contexts.

Another area of methodological inquiry emerges from registry data. Population based surveys that track residential choices over time suffer from well known problems of attrition. Even the vaunted PSID, over long observation intervals, exhibits substantial attrition. This attrition can be especially problematic for studies of residential mobility because it is often intrinsically bound up with the process of migration: Those who migrate frequently, or have loose ties to housing, may also be the most likely to drop out of a survey's sample. It is presently unclear how migration-related attrition might affect results from discrete choice models of residential mobility. Registry data have the advantage of providing nearly complete, longitudinal coverage of the population. Attrition is a negligible problem. Registry data provide an opportunity to construct "bad" samples, including samples with migration-related attrition, and compare results obtained with these samples to results obtained with either the full population, or proper, attrition-free random samples. This exercise can characterize the amount and sources of biases in discrete choice models estimated using samples

with attrition. They can also offer a test bed for techniques to assess the robustness of results obtained using “bad” samples.

Finally, the simulations models used in the present dissertation are quite simplistic. They abstract away from non-preference determinants of segregation, including income, education, home ownership, and social networks. Other simulation models have incorporated socioeconomic choice mechanisms (Bruch 2014; Fossett 2006a; Clark and Fossett 2008). These efforts allow for additional perspectives on how racial and non-racial factors interact in segregation processes, and also make it possible to characterize the evolution of segregation between agents sharing the same racial identification, but having disparate socioeconomic attainments. Adding socioeconomic factors back into the present agent based models would be a useful next step. In a more novel vein, no agent based models of segregation have yet included pathways of demographic reproduction, or tried to capture the effect of social networks on migration. Ideally, such efforts would be grounded in empirical work describing the effects, if any, of shared employment, schooling, or prior neighborhood co-residence on subsequent migration patterns. This work would finally extend current simulation models from shallow characterizations of social interaction, to a realm that allows thicker, and more human-like, domains of social interaction.

Whatever the promise of registry data, new methods, and simulation methods, this dissertation has broken new ground. It has exploited survey data in Los Angeles to show that while preferences may make important contributions to racial residential segregation, their contributions can be overstated. There appear to be other social forces intervening between preferences and actual neighborhood attainments that partially deflect people and families from attaining residence in neighborhoods that match their racial composition preferences. The forces in question make meaningful contributions to levels of residential segregation, over and above preferences. This suggests that there are still depths to plumb in accounting for extant residential segregation in the United States.

Tables

Table 2.1: Descriptives Statistics for Los Angeles Neighborhoods, 1998 and 2008

Variable	1998				2008			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Prop. Asian	0.13	0.14	0.00	0.82	0.15	0.15	0.00	0.86
Prop. Black	0.10	0.16	0.00	0.95	0.09	0.14	0.00	0.93
Prop. Latino	0.43	0.28	0.00	0.99	0.47	0.29	0.03	0.99
Prop. White	0.34	0.28	0.00	0.93	0.29	0.26	0.00	0.92
Spanish Language Density	0.40	0.17	0.00	0.83	0.34	0.17	0.00	0.88
Asian Language Density	0.29	0.17	0.00	0.92	0.27	0.21	0.00	1.00
Prop. Foreign Born	0.36	0.16	0.00	0.80	0.36	0.14	0.01	0.84
N Housing Units ^a	1.80	0.77	0.05	6.14	1.90	0.85	0.06	7.19
Median Household Income ^b	46.54	22.80	4.37	227.11	46.07	21.61	6.52	180.79
Prop. Age 25+ w/ Bachelor's	0.22	0.18	0.00	0.81	0.27	0.20	0.00	0.85
Prop. Owner Occupied Units	0.47	0.25	0.00	0.97	0.47	0.25	0.00	1.00

Note: N = 2,035 tracts with at least 50 housing units. Sources: 1990 and 2000 US Census, 2000 Neighborhood Change Database, 2005-2009 & 2006-2010 American Community Surveys.

^a Thousands of Units

^b Thousands of 1999 Dollars.

Table 2.2: Descriptive Statistics for L.A.FANS Respondents and Their Families

	Latino Rs		White Rs		Black Rs		Asian Rs	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age at Wave 2 Survey	44.208	13.951	54.777	15.963	45.054	16.622	51.714	15.412
Female	0.515	0.500	0.454	0.499	0.535	0.501	0.549	0.501
Married	0.497	0.500	0.599	0.491	0.258	0.439	0.596	0.494
Foreign Born	0.747	0.435	0.113	0.317	0.028	0.166	0.738	0.443
Bachelor's Degree or greater	0.063	0.242	0.516	0.501	0.354	0.480	0.482	0.503
W2 Family Income ^a	38.313	41.425	115.184	471.026	60.266	57.512	66.179	69.724
W1-W2 Avg. Family Income ^{a,b}	33.423	33.943	100.515	253.901	50.876	35.607	62.044	58.890
W2 Home owner	0.432	0.496	0.713	0.453	0.326	0.471	0.736	0.444
W1-W2 Home owner Average ^c	0.419	0.458	0.713	0.394	0.313	0.444	0.686	0.432
N Respondents	712		300		126		76	

Note: Statistics are weighted using L.A.FANS W2 cross-sectional weights for randomly selected adults.

^a Dollar values expressed in thousands of 1999 dollars and are inclusive of families and households with zero income.

^b based on average observed family income over the period for panel respondents, and the average of predicted and observed income for W2 only respondents.

^c based on average home ownership across W1 and W2 for panel respondents, and W2 report only for W2 only respondents.

Table 2.3: Descriptive Statistics for L.A.FANS Respondents' Stated Neighborhood Preferences

	Latino Rs		White Rs		Black Rs		Asian Rs	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A. Ideal SP								
Prop. Asian	0.144	0.132	0.164	0.133	0.184	0.085	0.361	0.195
Prop. Black	0.102	0.106	0.128	0.099	0.312	0.150	0.110	0.090
Prop. Latino	0.474	0.283	0.193	0.164	0.188	0.083	0.179	0.117
Prop. White	0.280	0.201	0.515	0.267	0.316	0.169	0.349	0.193
N Choices	712		300		126		76	
B. Rank SP								
1 st Ranked Choice								
Prop. Asian	0.081	0.129	0.119	0.144	0.075	0.094	0.358	0.179
Prop. Black	0.052	0.096	0.059	0.134	0.347	0.204	0.045	0.073
Prop. Latino	0.662	0.267	0.171	0.170	0.188	0.140	0.275	0.187
Prop. White	0.205	0.202	0.652	0.254	0.390	0.234	0.322	0.183
N Choices	712		300		126		76	
Full Choice Set								
Prop. Asian	0.072	0.123	0.105	0.140	0.065	0.128	0.346	0.226
Prop. Black	0.085	0.142	0.101	0.188	0.389	0.263	0.048	0.098
Prop. Latino	0.675	0.259	0.196	0.201	0.310	0.248	0.274	0.216
Prop. White	0.168	0.196	0.599	0.270	0.235	0.261	0.333	0.228
N Person-Alternatives	3560		1500		630		380	

Note: Statistics weighted using L.A.FANS Wave 2 cross-sectional weights for randomly selected adults.

Table 2.4: Descriptive Statistics for L.A.FANS Respondents' RH Neighborhoods

	Latino Rs		White Rs		Black Rs		Asian Rs	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A. Racial Composition								
Prop. Asian	0.098	0.115	0.144	0.133	0.129	0.125	0.288	0.129
Prop. Black	0.066	0.094	0.057	0.052	0.162	0.126	0.059	0.041
Prop. Latino	0.681	0.232	0.279	0.229	0.416	0.234	0.374	0.212
Prop. White	0.151	0.158	0.513	0.250	0.287	0.252	0.271	0.147
B. Socio-Economic Characteristics								
Median Household Income ^a	35.668	12.371	60.389	25.178	39.675	15.842	47.601	15.426
R's Position in Income Distribution	0.406	0.264	0.552	0.291	0.542	0.257	0.548	0.289
Prop. w/ Bachelor's Degree	0.125	0.105	0.363	0.217	0.250	0.200	0.271	0.116
Prop. Owner occupied	0.416	0.220	0.600	0.234	0.393	0.225	0.521	0.242
C. Residential Mobility								
Remained in Origin House	0.979	0.144	0.986	0.117	0.977	0.149	0.985	0.121
Approx. Yearly Mobility Rate	0.082		0.055		0.088		0.059	
N Person-Quarters	22,344		9,465		3,963		2,364	
N Respondents	712		300		126		76	
N Moves	560		156		129		53	
N Rs with at least 1 Move	376		119		71		34	

Note: Statistics are weighted using L.A.FANS W2 cross-sectional weights for randomly selected adults.
^a CPI adjusted to thousands of 1999 dollars.

Table 2.5: Racial Composition Coefficients from Joint SP and RH Conditional Logistic Regression Models of Neighborhood Choice

Variable	Latino Rs		White Rs		Black Rs		Asian Rs	
	b	b/se	b	b/se	b	b/se	b	b/se
<i>Main Racial Composition Effects (RH Scenarios)</i>								
Prop. Asian	-2.535	-1.22	-8.137	-2.36	-0.518	-0.09	2.726	0.50
Prop. Asian ²	-2.258	-0.84	2.659	0.74	-5.185	-0.88	-8.335	-1.37
Prop. Black	-2.140	-1.57	-1.005	-0.21	4.429	1.40	-12.170	-2.64
Prop. Black ²	-3.547	-1.65	-13.806	-1.05	-9.994	-2.45	8.771	2.00
Prop. Latino	-0.910	-0.31	-7.896	-2.23	-4.514	-0.73	-4.130	-0.80
Prop. Latino ²	-0.918	-0.55	1.483	0.65	0.521	0.15	0.172	0.05
Prop. White ²	-6.450	-3.24	-4.224	-1.92	-4.963	-1.27	-3.977	-1.18
<i>Own House × ...</i>								
Prop. Asian	7.755	3.01	15.384	2.79	-0.105	-0.01	-7.038	-1.39
Prop. Asian ²	-3.225	-0.67	-13.024	-2.17	14.702	1.01	-0.658	-0.06
Prop. Black	4.563	2.74	-1.180	-0.27	0.693	0.14	2.060	0.32
Prop. Black ²	0.603	0.22	16.661	1.62	4.111	0.90	-5.887	-0.62
Prop. Latino	-0.123	-0.04	8.618	1.75	2.782	0.50	-15.664	-2.02
Prop. Latino ²	3.886	1.84	-0.927	-0.29	2.717	0.75	7.909	1.81
Prop. White ²	6.827	2.65	6.808	2.35	7.036	1.37	-12.389	-2.17
<i>Ideal SP Scenario × ...</i>								
Prop. Asian	-0.970	-0.39	19.799	3.20	18.742	2.09	-0.554	-0.08
Prop. Asian ²	-9.855	-2.48	-23.093	-2.29	-69.827	-4.31	1.574	0.21
Prop. Black	-2.902	-1.71	17.328	3.00	-3.770	-0.47	12.924	2.51
Prop. Black ²	-9.751	-2.99	-27.555	-1.94	-3.878	-0.63	-30.481	-3.34
Prop. Latino	-8.134	-2.38	17.839	3.31	21.962	2.41	5.368	0.75
Prop. Latino ²	6.696	3.49	-16.630	-2.41	-62.627	-5.04	-11.139	-1.39
Prop. White ²	-2.618	-0.99	10.445	3.72	-11.641	-1.10	1.104	0.21
<i>Rank SP Scenario × ...</i>								
Prop. Asian	2.182	0.99	8.938	2.50	0.098	0.02	-1.055	-0.18
Prop. Asian ²	-0.388	-0.13	-10.147	-2.62	-2.703	-0.44	4.539	0.73
Prop. Black	-2.296	-1.48	-2.759	-0.56	-5.606	-1.61	6.500	1.21
Prop. Black ²	3.984	1.67	10.890	0.82	7.441	1.74	-5.075	-0.97
Prop. Latino	0.854	0.28	6.973	1.86	2.003	0.31	-0.317	-0.06
Prop. Latino ²	0.168	0.10	-5.012	-2.11	-3.718	-1.03	1.705	0.51
Prop. White ²	4.985	2.36	2.425	1.04	-0.414	-0.10	1.927	0.53
<i>Continued</i>								

Table 2.5 Continued: Racial Composition Coefficients from Joint SP+RH Models

Variable	Latino Rs		White Rs		Black Rs		Asian Rs	
	b	b/se	b	b/se	b	b/se	b	b/se
N Person-Sit.-Alts.	2,273,940		965,700		404,364		241,264	
N Person-Sits.	25,904		10,965		4,593		2,744	
N Moves in RH Sits.	560		156		129		53	
N Respondents	712		300		126		76	
Estimated Parameters	50		50		50		50	
log-likelihood	-12,609		-4,524		-2,410		-1,266	

Note: Models estimated using Manski-Lerman Weights. Models also included a correction for a sampling of alternatives. Correction is equal to -1 times the log of the sampling fraction.

Source: L.A.FANS Waves 1 and 2.

Table 2.6: Socioeconomic, Demographic, and Proximity Coefficients from Joint SP and RH Conditional Logistic Regression Models of Neighborhood Choice

Variable	Latino Rs		White Rs		Black Rs		Asian Rs	
	b	b/se	b	b/se	b	b/se	b	b/se
<i>Neighborhood Size</i>								
Log Housing Units	0.815	5.60	0.425	2.39	0.691	2.82	1.343	3.33
<i>Proximity</i>								
Own House	7.575	3.77	1.490	0.47	6.158	1.52	20.773	3.92
Own Neighborhood	—	—	—	—	—	—	—	—
1st Order Adjacent	-0.570	-2.53	-0.611	-1.88	0.182	0.32	-0.154	-0.25
2nd-4th Order Adjacent	-1.860	-6.73	-1.502	-4.07	-0.141	-0.28	-0.996	-1.65
5 th + Order Adjacent	-3.467	-7.63	-3.303	-5.57	-0.280	-0.35	-2.033	-2.10
Sqrt(Distance to Tract)	-0.748	-5.77	-0.635	-4.81	-1.298	-6.62	-1.351	-5.29
<i>Income Effects</i>								
Income Position	2.770	2.13	5.293	2.45	3.300	1.15	7.625	2.02
Income Position ²	-4.500	-3.73	-3.090	-1.90	-3.018	-1.38	-6.581	-1.86
<i>Own House × ...</i>								
Income Position	-1.573	-1.65	-0.830	-0.53	1.598	0.72	1.146	0.33
Income Position ²	2.034	1.92	0.496	0.38	-1.235	-0.54	-0.334	-0.11
<i>Education Effects</i>								
Prop. w/ Bachelor's	-1.235	-1.27	-1.375	-1.40	-1.285	-0.74	-8.328	-4.00
<i>Has Bachelor's × ...</i>								
Prop. w/ Bachelor's	0.303	0.23	2.890	4.02	3.550	3.12	4.869	3.15
<i>Home ownership Effects</i>								
Prop. Owner Occupied	-0.418	-1.16	-0.382	-0.62	-0.653	-0.79	-1.594	-1.37
<i>Home owner × ...</i>								
Prop. Owner Occupied	3.232	6.14	2.878	4.23	3.414	2.90	3.742	2.71
<i>Selection into Migration</i>								
<i>Own House × ...</i>								
Age	0.025	4.74	0.038	5.23	0.039	3.19	0.052	2.96
Has Bachelor's Degree	0.039	0.15	0.028	0.14	-0.298	-0.80	0.449	1.58
Home Owner	0.648	4.18	0.469	1.90	0.883	2.14	1.005	2.14
Married	0.042	0.38	0.169	0.86	-0.476	-1.35	-0.684	-1.43
<i>Language</i>								
Spanish Lang. Density	-1.240	-0.88	2.743	1.77	3.681	1.25	-7.660	-2.15
Spanish Lang. Density ²	0.272	0.19	-2.663	-1.14	-4.610	-1.36	3.826	0.90
Asian Lang. Density	-1.050	-1.80	0.523	0.39	-0.181	-0.12	6.469	2.08
<i>Continued</i>								

Table 2.6 Continued: Non-Racial Coefficients from Joint SP+RH Models

Variable	Latino Rs		White Rs		Black Rs		Asian Rs	
	b	b/se	b	b/se	b	b/se	b	b/se
Asian Lang. Density ²	0.749	1.05	-0.566	-0.30	-0.711	-0.40	-9.340	-2.05
N Person-Sit.-Alts.	2,273,940		965,700		404,364		241,264	
N Person-Sits.	25,904		10,965		4,593		2,744	
N Moves in RH Sits.	560		156		129		53	
N Respondents	712		300		126		76	
Estimated Parameters	50		50		50		50	
log-likelihood	-12,609		-4,524		-2,410		-1,266	

Note: Models estimated using Manski-Lerman Weights. Models also included a correction for a sampling of alternatives. Correction is equal to -1 times the log of the sampling fraction.

Source: L.A.FANS Waves 1 and 2.

Table 2.7: Wald Tests of Racial Composition Effects Within RH and SP Neighborhood Choice Scenarios

Racial Composition ^a	(1) RH Scenario			(2) Ideal SP Scenario			(3) Rank SP Scenario		
	χ^2	<i>df</i>	<i>p</i>	χ^2	<i>df</i>	<i>p</i>	χ^2	<i>df</i>	<i>p</i>
A. Latino Rs (N=712)									
Prop. Asian	7.53	2	0.023	67.22	2	<0.001	10.04	2	0.007
Prop. Black	10.24	2	0.006	85.48	2	<0.001	86.92	2	<0.001
Prop. Latino	5.05	2	0.080	43.65	2	<0.001	11.16	2	0.0038
Prop. White	10.50	1	0.001	29.51	1	<0.001	4.83	1	0.028
All Terms	52.35	7	<0.001	987.41	7	<0.001	196.77	7	<0.001
B. White Rs (N=300)									
Prop. Asian	8.19	2	0.017	4.98	2	0.083	30.11	2	<0.001
Prop. Black	8.10	2	0.017	80.47	2	<0.001	38.38	2	<0.001
Prop. Latino	14.42	2	0.001	4.97	2	0.084	35.96	2	<0.001
Prop. White	3.70	1	0.055	10.44	1	0.001	4.47	1	0.0344
All Terms	37.85	7	<0.001	362.29	7	<0.001	131.79	7	<0.001
C. Black Rs (N=126)									
Prop. Asian	2.96	2	0.228	25.71	2	<0.001	18.27	2	<0.001
Prop. Black	6.03	2	0.049	20.67	2	<0.001	23.43	2	<0.001
Prop. Latino	2.57	2	0.277	29.44	2	<0.001	50.65	2	<0.001
Prop. White	1.60	1	0.206	2.98	1	0.084	26.55	1	<0.001
All Terms	16.68	7	0.020	56.10	7	<0.001	69.12	69.1	<0.001
D. Asian Rs (N=76)									
Prop. Asian	2.43	2	0.296	3.09	2	0.213	6.68	2	0.035
Prop. Black	7.00	2	0.030	11.79	2	0.003	5.42	2	0.067
Prop. Latino	2.07	2	0.356	4.83	2	0.090	4.99	2	0.082
Prop. White	1.39	1	0.238	0.56	1	0.454	1.84	1	0.175
All Terms	22.71	7	0.002	32.37	7	<0.001	21.15	7	0.004

Note: RH portion of models include all socioeconomic and non-racial controls.

^a Racial composition terms include second order polynomials in proportion Asian, Black, Latino, and White, omitting the linear proportion White term.

Table 2.8: Wald Tests of Differences in Racial Composition Effects: SP vs. RH Neighborhood Choice Scenarios

Racial Composition ^a	(1) Ideal SP vs. RH			(2) Rank SP vs. RH		
	χ^2	<i>df</i>	<i>p</i>	χ^2	<i>df</i>	<i>p</i>
A. Latino Rs (N=712)						
Prop. Asian	12.63	2	0.002	1.59	2	0.451
Prop. Black	18.88	2	<0.001	3.25	2	0.197
Prop. Latino	15.77	2	<0.001	0.92	2	0.632
Prop. White	0.98	1	0.322	5.55	1	0.018
All Terms	129.98	7	<0.001	39.94	7	<0.001
B. White Rs (N=300)						
Prop. Asian	12.52	2	0.002	7.54	2	0.023
Prop. Black	11.30	2	0.004	0.81	2	0.666
Prop. Latino	12.15	2	0.002	4.43	2	0.109
Prop. White	13.82	1	<0.001	1.08	1	0.299
All Terms	15.74	7	0.028	14.72	7	0.040
C. Black Rs (N=126)						
Prop. Asian	19.72	2	<0.001	0.53	2	0.767
Prop. Black	1.96	2	0.375	3.60	2	0.166
Prop. Latino	26.48	2	<0.001	3.90	2	0.142
Prop. White	1.21	1	0.272	0.01	1	0.919
All Terms	38.64	7	<0.001	12.19	7	0.094
D. Asian Rs (N=76)						
Prop. Asian	0.05	2	0.976	0.80	2	0.670
Prop. Black	11.51	2	0.003	1.47	2	0.480
Prop. Latino	2.12	2	0.346	0.72	2	0.696
Prop. White	0.04	1	0.837	0.28	1	0.598
All Terms	16.57	7	0.020	5.68	7	0.578

Note: RH portion of models include all socioeconomic and other non-racial controls.

^a Racial composition terms include second order polynomials in proportion Asian, Black, Latino, and White, omitting the linear proportion White.

Table 2.9: Cross Scenario Dissimilarities in Predicted Probabilities from Joint SP-RH Models of Neighborhood Choice, Los Angeles Neighborhood Choice Set

Model/Scenario		Model/Scenario				
		(1)	(2)	(3A)	(3B)	(4)
A. Latinos						
(1)	SP - Ideal	0.00				
(2)	SP - Ranked	0.41	0.00			
(3A)	RH - Race Only Model ^a	0.25	0.27	0.00		
(3B)	RH - Full Model ^b	0.30	0.24	0.07	0.00	
(4)	Indifference ^c	0.48	0.13	0.33	0.28	0.00
B. Whites						
(1)	SP - Ideal	0.00				
(2)	SP - Ranked	0.30	0.00			
(3A)	RH - Race Only Model ^a	0.23	0.13	0.00		
(3B)	RH - Full Model ^b	0.24	0.14	0.08	0.00	
(4)	Indifference ^c	0.55	0.34	0.37	0.36	0.00
C. Blacks						
(1)	SP - Ideal	0.00				
(2)	SP - Ranked	0.60	0.00			
(3A)	RH - Race Only Model ^a	0.64	0.34	0.00		
(3B)	RH - Full Model ^b	0.63	0.16	0.20	0.00	
(4)	Indifference ^c	0.75	0.26	0.39	0.24	0.00

Note: Predicted probabilities derived from models that include second order polynomials in racial composition, omitting the linear proportion White term. Based on a choice set of 2,035 Los Angeles County Neighborhoods with at least 50 housing units. Racial composition averaged over the 1998-2008 period. Non-racial characteristics fixed at constant values.

^a Includes racial composition terms and adjustment for sampling of alternatives.

^b As in (a), but including all non-racial controls discussed in the text.

^c Equal probability of choosing each neighborhood in the choice set.

Table 2.10: Predicted Neighborhood Racial Composition based on a Los Angeles Neighborhood Choice Set, from Conditional Logit Models of SP and RH Data

Model/Scenario	% Asian	% Black	% Latino	% White
A. Latino Rs				
SP - Ideal	6.5	3.9	75.1	14.5
SP - Ranked	13.3	5.7	48.6	32.4
RH - Race Only Model	11.1	6.8	64.5	17.6
RH - Full Model	11.5	8.3	61.5	18.8
B. White Rs				
SP - Ideal	12.2	4.9	17.0	65.9
SP - Ranked	15.1	5.5	28.3	51.1
RH - Race Only Model	14.8	5.5	25.9	53.8
RH - Full Model	13.0	5.7	26.9	54.4
C. Black Rs				
SP - Ideal	16.8	18.0	24.8	40.4
SP - Ranked	15.1	12.6	36.1	36.2
RH - Race Only Model	12.3	18.3	44.7	24.7
RH - Full Model	14.4	13.9	41.5	30.3
D. Indifferent Housing Seeker				
All Scenarios	13.6	9.5	44.8	32.2

Predictions produced using a choice set of 2,035 Los Angeles Neighborhoods, with racial composition averaged over the 1998-2008 period. All non-racial variables held constant.

Ideal SP, rank SP, and race only RH models include only second order polynomials in the proportion Asian, Black, Latino, and White, excluding the linear proportion White term, as explanatory variables. The full RH model includes all other non-racial controls discussed in the text.

Table 3.1: Racial Composition Coefficients from Conditional Logit Models of Whites' Rank SP

Variable	Homogeneous Effects ^a		Heterogeneous Effects by Rank ^b							
	All Ranks		Rank 1		Rank 2		Rank 3		Rank 4	
	b	b/se	b	b/se	b	b/se	b	b/se	b	b/se
Prop. Asian	0.80	0.60	8.11	2.74	-2.33	-0.87	0.72	0.25	-6.09	-2.21
Prop. Asian Sq.	-7.49	-4.45	-13.27	-3.28	-5.39	-1.65	-8.13	-2.24	-2.26	-0.59
Prop. Black	-3.76	-2.57	3.50	1.26	-8.49	-3.02	-0.25	-0.07	-9.75	-3.15
Prop. Black Sq.	-2.92	-1.92	-6.21	-2.01	2.01	0.69	-16.73	-3.18	1.56	0.59
Prop. Latino	-0.92	-0.65	3.92	1.40	-4.49	-1.69	-1.53	-0.52	-1.70	-0.63
Prop. Latino Sq.	-3.53	-3.32	-4.11	-1.94	-2.69	-1.31	-3.46	-1.78	-5.97	-3.20
Prop. White Sq.	-1.80	-2.12	1.99	1.28	-3.77	-2.40	-2.01	-1.11	-4.95	-2.78
N	4200		4200							
N Choices	1200		1200							
N Respondents	300		300							
Parameters	7		28							
log-likelihood	-1283.71		-1258.80							
BIC: N = 300	2607.34		2677.31							
BIC: N = 1200	2617.05		2716.13							

Source: L.A.FANS Wave 2

^a Uniform racial composition effects across ranks, i.e., no differences in effects across rankings.

^b Model includes separate effects for each rank with no omitted category.

Table 3.2: Racial Composition Coefficients from Conditional Logistic and Mixed Logit Models of Whites' Rank SP

Variable	Mixed Logit					
	Conditional Logit		Univariate Mixing ^a		Multivariate Mixing ^b	
	b	b/se	b	b/se	b	b/se
Coefficients						
Prop. Asian	0.801	0.60	2.107	0.89	0.658	0.20
Prop. Asian Sq.	-7.488	-4.45	-11.412	-3.98	-14.862	-4.01
Prop. Black	-3.764	-2.57	-3.516	-1.36	-6.856	-1.75
Prop. Black Sq.	-2.917	-1.92	-12.820	-3.28	-18.320	-3.02
Prop. Latino	-0.923	-0.65	0.613	0.25	-1.698	-0.51
Prop. Latino Sq.	-3.529	-3.32	-9.482	-4.51	-10.632	-3.65
Prop. White Sq.	-1.799	-2.12	-1.926	-1.38	-3.651	-1.85
Mix Logit Standard Deviations						
Prop. Asian			3.868	3.63	8.581	2.00
Prop. Asian Sq.			1.107	0.46	11.162	2.85
Prop. Black			6.978	5.80	12.508	3.43
Prop. Black Sq.			7.372	2.07	13.443	3.87
Prop. Latino			1.916	1.55	19.567	5.33
Prop. Latino Sq.			4.602	4.11	29.319	4.22
Prop. White Sq.			4.690	8.01	10.404	4.70
N	4200		4200		4200	
N Choices	1200		1200		1200	
N Respondents	300		300		300	
Model df	7		14		35	
log-likelihood	-1283.71		-1212.11		-1183.94	
BIC: N = 300	2607.34		2504.08		2567.51	
BIC: N = 1200	2617.05		2523.48		2616.03	

Source: L.A.FANS Wave 2

^a Independent, univariate normal mixing distributions estimated for each racial composition effect.

^b Multivariate normal mixing distribution allows correlations between preference components within respondents. See Table 3.3 for covariances of random effects.

Table 3.3: Estimated Covariances of Racial Composition Coefficients for Whites, Mixed Logit Models of Rank SP Data

	Prop. Asian	Prop. Asian ²	Prop. Black	Prop. Black ²	Prop. Latino	Prop. Latino ²	Prop. White ²
Prop. Asian	73.64 (73.68)						
Prop. Asian ²	40.36 (28.48)	124.59 (87.44)					
Prop. Black	25.19 (54.54)	89.64 (60.00)	382.87** (143.66)				
Prop. Black ²	171.88* (87.28)	143.87 (117.15)	-269.52 (189.83)	859.58* (407.33)			
Prop. Latino	46.23 (62.87)	64.30 (57.57)	198.31* (85.27)	-60.33 (129.69)	156.44† (91.27)		
Prop. Latino ²	66.17† (36.39)	64.08 (65.26)	-138.44† (78.97)	377.86* (168.81)	-31.0 (59.37)	180.71† (93.45)	
Prop. White ²	51.81 (47.98)	92.65† (50.61)	36.25 (51.66)	165.95* (76.14)	65.16 (55.48)	83.87* (35.80)	108.25* (46.04)

Source: L.A.FANS Wave 2.

Note: Standard errors in parentheses.

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .1$

Table 3.4: Racial Composition Coefficients from Conditional and Mixed Logit Models of Whites' Neighborhood Choices in Rank SP and Ideal SP Vignettes

Variable	Separate Ideal SP & Rank SP Effects				Main Effect + Ideal SP Interactions ^a			
	clogit		mixlogit		clogit		mixlogit ^b	
	b	b/se	b	b/se	b	b/se	b	b/se
Racial Composition Effects								
<i>Rank SP</i>								
Prop. Asian	-0.388	-1.23	-0.580	-0.94	-0.388	-1.23	-0.882	-2.31
Prop. Black	-3.025	-10.04	-6.992	-7.75	-3.025	-10.04	-4.480	-9.98
Prop. Latino	-1.396	-5.95	-2.977	-5.42	-1.396	-5.95	-2.260	-6.93
<i>Ideal SP</i>								
Prop. Asian	-3.722	-13.31	-3.866	-10.01	-3.334	-8.97	-3.028	-5.91
Prop. Black	-4.947	-16.95	-5.486	-11.18	-1.922	-6.37	-1.417	-2.68
Prop. Latino	-3.544	-13.16	-3.755	-9.72	-2.148	-7.38	-1.527	-3.50
Standard Deviations								
<i>Rank SP</i>								
Prop. Asian			6.647	7.23			3.200	7.89
Prop. Black			9.401	7.85			4.677	9.86
Prop. Latino			6.192	7.37			3.343	9.07
<i>Ideal SP</i>								
Prop. Asian			2.476	5.32			3.200	7.89
Prop. Black			3.364	6.21			4.677	9.86
Prop. Latino			2.656	5.64			3.343	9.07
N	19200		19200		19200		19200	
N Choices	1500		1500		1500		1500	
N Rs	300		300		300		300	
Parameters	6		27		6		12	
log-likelihood	-3116.688		-3008.79		-3116.688		-3037.62	
BIC: N = 1500	6277.256		6215.045		6277.256		6162.996	
BIC: N = 300	6267.599		6171.591		6267.599		6143.682	

Source: L.A.FANS Wave 2

Note: Models include a correction for the sampling of alternatives in the ideal SP case.

^a Includes main racial composition effects common to rank and ideal SP scenarios, and an interaction between racial composition and a dummy identifying the ideal SP scenario.

^b In this specification, random effect standard deviations are identical in ideal SP and rank SP cases.

Table 3.5: Estimated Covariances of Racial Composition Coefficients for Whites, Mixed Logit Models of Rank and Ideal SP Data

Rank Stated Preferences			Ideal Stated Preferences			
	Prop. Asian	Prop. Black	Prop. Latino	Prop. Asian	Prop. Black	Prop. Latino
Rank Stated Preferences						
Prop. Asian	44.188*** (12.216)					
Prop. Black	39.049** (11.687)	88.378*** (22.508)				
Prop. Latino	21.946** (6.942)	40.128*** (9.725)	38.342*** (10.408)			
Ideal Stated Preferences						
Prop. Asian	12.448*** (3.473)	22.428*** (5.334)	11.496*** (3.133)	6.128** (2.304)		
Prop. Black	14.960*** (4.076)	30.725*** (6.682)	16.628*** (3.924)	8.239** (2.452)	11.314** (3.644)	
Prop. Latino	11.269*** (3.158)	24.205*** (5.489)	13.355*** (3.345)	6.454** (2.020)	8.924** (2.597)	7.055** (2.504)

Source: L.A.FANS Wave 2.

Note: Standard errors in parentheses.

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .1$

Table 3.6: Racial Composition Coefficients from Conditional and Mixed Logit Models of Whites' Neighborhood Choices in RH and Rank SP Vignettes

Variable	Separate RH & Rank SP Effects				Main Effect + Rank SP Interactions ^a			
	clogit		mixlogit		clogit		mixlogit ^b	
	b	b/se	b	b/se	b	b/se	b	b/se
Racial Composition Effects								
<i>RH</i>								
Prop. Asian	-1.382	-2.42	-1.464	-2.32	-1.382	-2.42	-1.835	-2.77
Prop. Black	-3.667	-4.25	-3.963	-3.49	-3.667	-4.25	-5.442	-5.03
Prop. Latino	-1.816	-3.40	-1.947	-3.11	-1.816	-3.40	-1.798	-2.83
<i>Rank SP</i>								
Prop. Asian	-0.388	-1.23	-0.309	-0.51	0.995	1.63	1.436	1.95
Prop. Black	-3.025	-10.04	-6.349	-7.56	0.642	0.71	1.450	1.40
Prop. Latino	-1.396	-5.95	-2.614	-5.08	0.420	0.71	-0.032	-0.05
Standard Deviations								
<i>RH</i>								
Prop. Asian			1.329	1.70			2.875	6.31
Prop. Black			1.737	1.56			4.018	7.40
Prop. Latino			0.427	1.09			1.361	3.92
<i>Rank SP</i>								
Prop. Asian			6.155	6.84			2.875	6.31
Prop. Black			8.428	7.80			4.018	7.40
Prop. Latino			5.762	6.81			1.361	3.92
N	950700		950700		950700		950700	
N Choices	10965		10965		10965		10965	
N Respondents	300		300		300		300	
Parameters	22		43		22		28	
log-likelihood	-2895.89		-2826.15		-2895.89		-2867.80	
BIC: N = 300	5917.263		5897.566		5917.263		5895.304	
BIC: N = 10965	5996.43		6052.31		5996.43		5996.07	

Source: L.A.FANS Wave 2

Note: Models include a correction for the sampling of alternatives in the RH case.

RH portions of models also include controls for neighborhood proximity, log housing units, and matching on income, education, and home ownership.

^a Includes main racial composition effects common to RH and rank SP scenarios, and an interaction between racial composition and a dummy identifying the rank SP scenario.

^b Random effect standard deviations are identical in RH and rank SP cases in this specification.

Table 3.7: Estimated Covariances of Racial Composition Coefficients for Whites, Mixed Logit Models of Rank SP and RH Data

	Residential History			Rank Stated Preferences		
	Prop. Asian	Prop. Black	Prop. Latino	Prop. Asian	Prop. Black	Prop. Latino
Residential History						
Prop. Asian	1.767 (2.074)					
Prop. Black	2.286 (2.211)	3.017 (3.877)				
Prop. Latino	-0.200 (0.628)	-0.290 (0.899)	0.183 (0.336)			
Rank Stated Preferences						
Prop. Asian	2.219 (4.918)	3.946 (7.517)	-2.001 (2.667)	37.889** (11.078)		
Prop. Black	6.003 (5.703)	8.050 (9.050)	-3.482 (3.415)	33.477*** (8.838)	71.036*** (18.220)	
Prop. Latino	-2.217 (4.415)	-2.572 (6.441)	-1.938 (2.280)	18.502** (6.068)	31.661*** (9.079)	33.204** (9.748)

Source: L.A.FANS Wave 2.

Note: Standard errors in parentheses.

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .1$

Table 4.1: Entropy Index by Scale and Source of Racial Composition Coefficients, from Schelling Segregation Models

	$\theta = 1$			$\theta = 2$			$\theta = 5$		
	SP-I	SP-R	RH	SP-I	SP-R	RH	SP-I	SP-R	RH
Entropy Index	0.739 (0.017)	0.092 (0.003)	0.498 (0.012)	0.738 (0.013)	0.324 (0.014)	0.781 (0.012)	0.724 (0.009)	0.660 (0.020)	0.839 (0.005)
<i>Decomposition of Entropy Index by Group</i>									
Latino	0.215 (0.004)	0.021 (0.001)	0.125 (0.004)	0.217 (0.006)	0.097 (0.007)	0.215 (0.004)	0.220 (0.007)	0.208 (0.008)	0.250 (0.003)
White	0.194 (0.006)	0.019 (0.001)	0.127 (0.007)	0.187 (0.006)	0.079 (0.005)	0.232 (0.004)	0.182 (0.005)	0.174 (0.011)	0.238 (0.005)
Black	0.165 (0.004)	0.030 (0.002)	0.053 (0.003)	0.158 (0.003)	0.096 (0.004)	0.125 (0.006)	0.152 (0.004)	0.137 (0.004)	0.147 (0.004)
Asian	0.165 (0.008)	0.022 (0.001)	0.192 (0.002)	0.175 (0.007)	0.051 (0.003)	0.208 (0.002)	0.169 (0.003)	0.141 (0.006)	0.204 (0.003)

Note: Standard deviations in parentheses.

Statistics pertain to state of simulations after 10,000 model time units, taken over 10 randomly seeded simulation runs for each set of racial composition coefficients and each value of the scale parameter, θ .

Ideal SP, rank SP, and RH simulations are based on racial composition coefficients derived from discrete choice models of preferences and inter-neighborhood migration in L.A.FANS presented in Table 2.5.

Table 4.2: Pairwise Indices of Dissimilarity by Scale and Source of Racial Composition Coefficients, Schelling Residential Segregation Models

	$\theta = 1$			$\theta = 2$			$\theta = 5$		
	SP-I	SP-R	RH	SP-I	SP-R	RH	SP-I	SP-R	RH
Latino-White	0.881 (0.012)	0.265 (0.006)	0.642 (0.021)	0.873 (0.020)	0.560 (0.019)	0.939 (0.006)	0.890 (0.018)	0.831 (0.021)	0.970 (0.008)
Latino-Black	0.894 (0.010)	0.366 (0.012)	0.380 (0.016)	0.888 (0.013)	0.760 (0.021)	0.699 (0.025)	0.876 (0.009)	0.890 (0.024)	0.856 (0.019)
Latino-Asian	0.930 (0.011)	0.317 (0.008)	0.933 (0.006)	0.952 (0.014)	0.585 (0.017)	0.995 (0.002)	0.953 (0.011)	0.925 (0.012)	0.999 (0.001)
Black-White	0.832 (0.013)	0.344 (0.013)	0.590 (0.026)	0.812 (0.015)	0.615 (0.019)	0.880 (0.028)	0.801 (0.018)	0.787 (0.026)	0.852 (0.030)
Asian-White	0.821 (0.024)	0.303 (0.012)	0.943 (0.006)	0.848 (0.030)	0.518 (0.016)	0.995 (0.001)	0.819 (0.015)	0.807 (0.036)	0.999 (0.001)
Black-Asian	0.898 (0.033)	0.384 (0.011)	0.911 (0.017)	0.890 (0.022)	0.536 (0.018)	0.911 (0.023)	0.874 (0.026)	0.689 (0.020)	0.853 (0.027)

Note: Standard deviations in parentheses.

Statistics pertain to state of simulations after 10,000 model time units, taken over 10 randomly seeded simulation runs for each set of racial composition coefficients and each value of the scale parameter, θ .

Ideal SP, rank SP, and RH simulations are based on racial composition coefficients derived from discrete choice models of preferences and inter-neighborhood migration in L.A.FANS presented in Table 2.5.

Figures

Figure 2.1: Sample Ranked Neighborhood Vignette Card

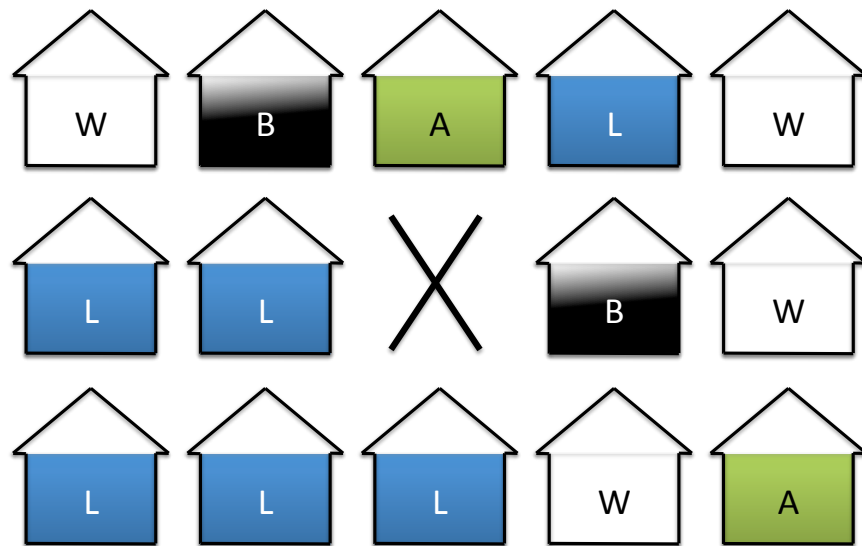


Figure 2.2: Ideal Neighborhood Vignette Card

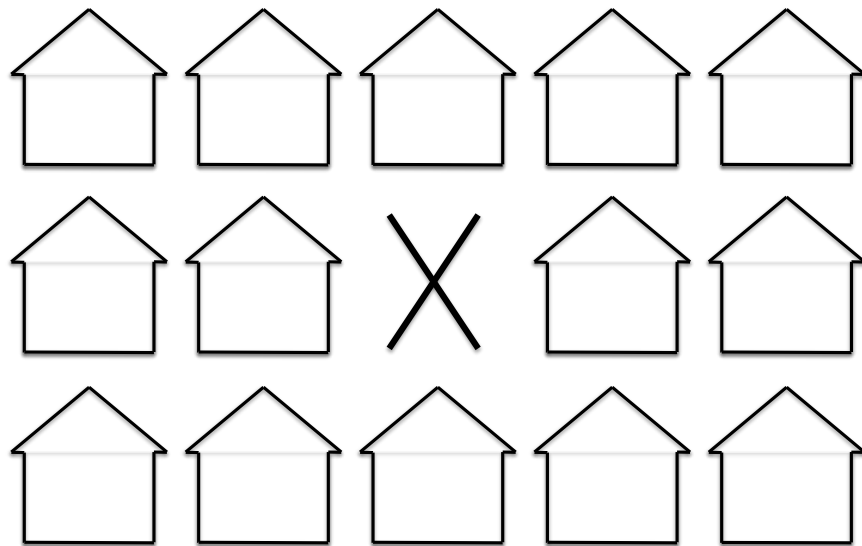


Figure 2.3: Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino

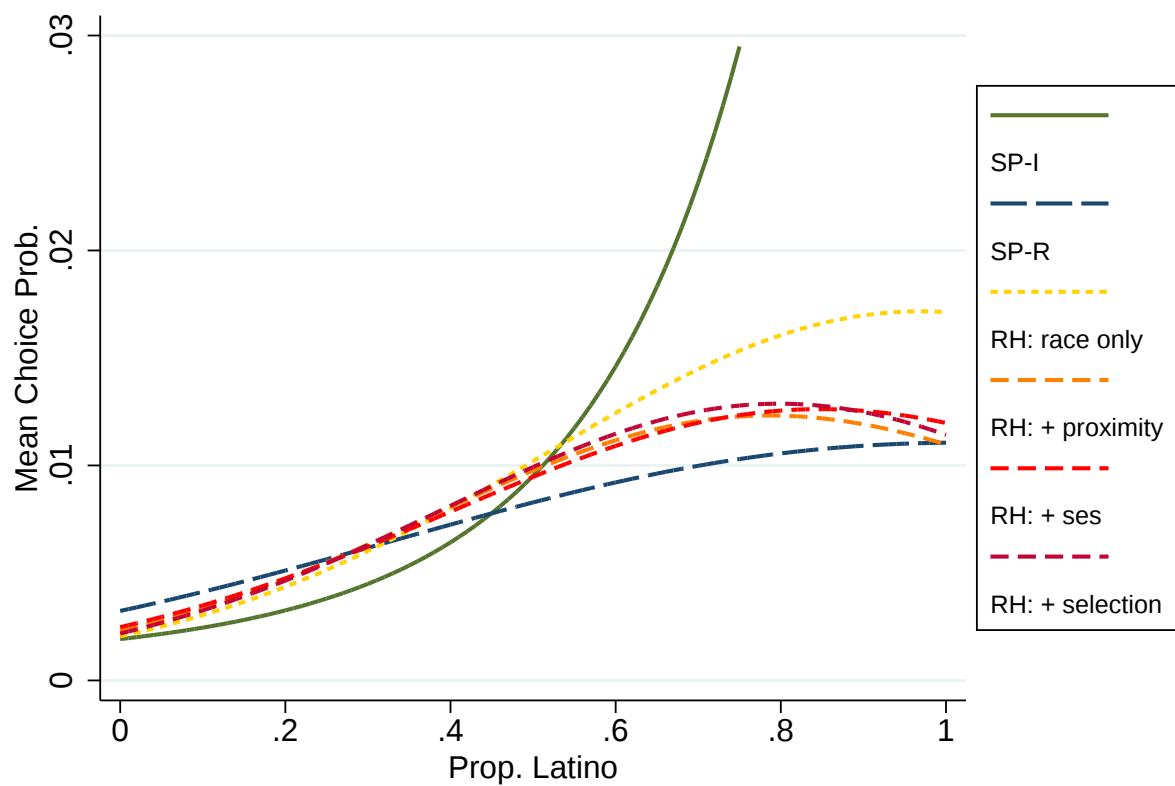


Figure 2.4: Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black

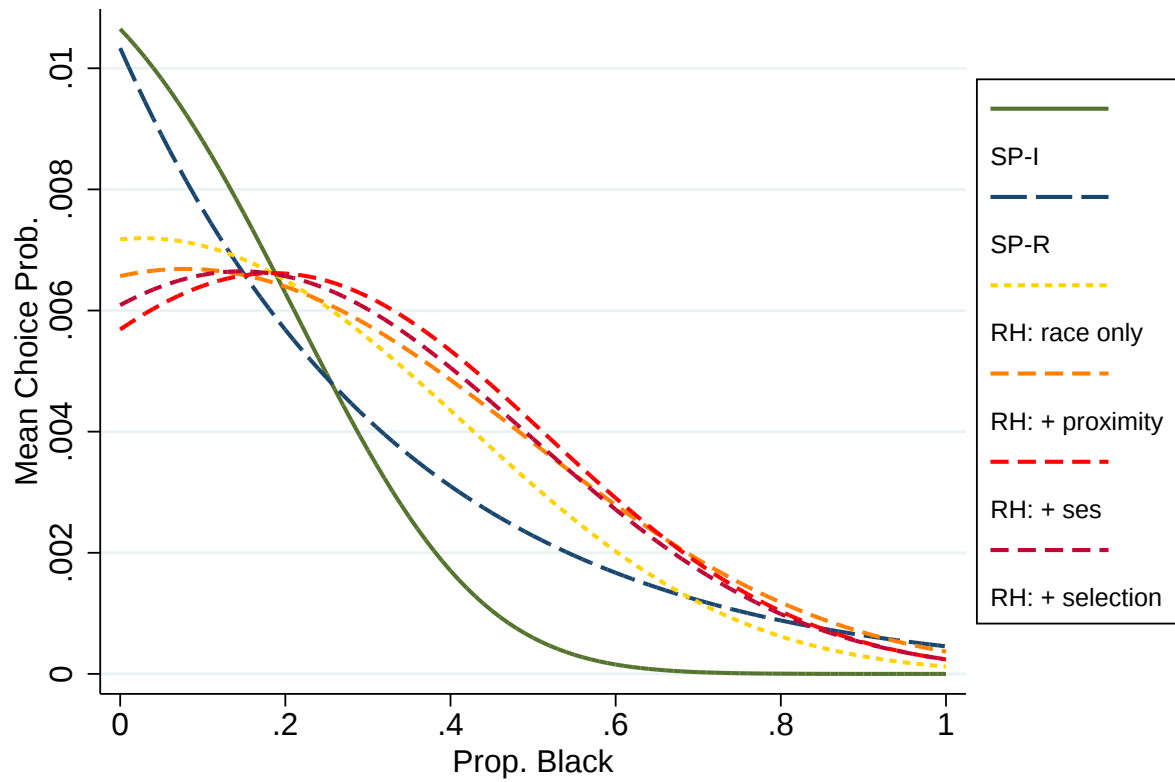


Figure 2.5: Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White

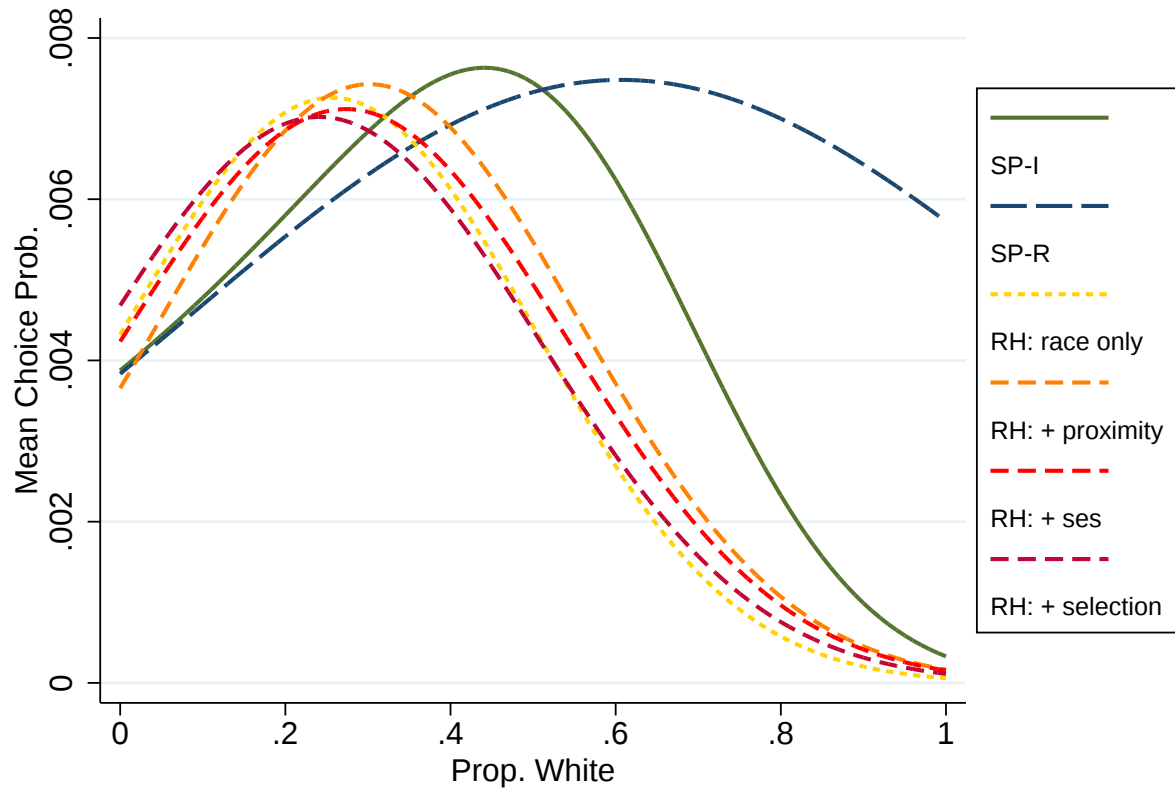


Figure 2.6: Latino Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian

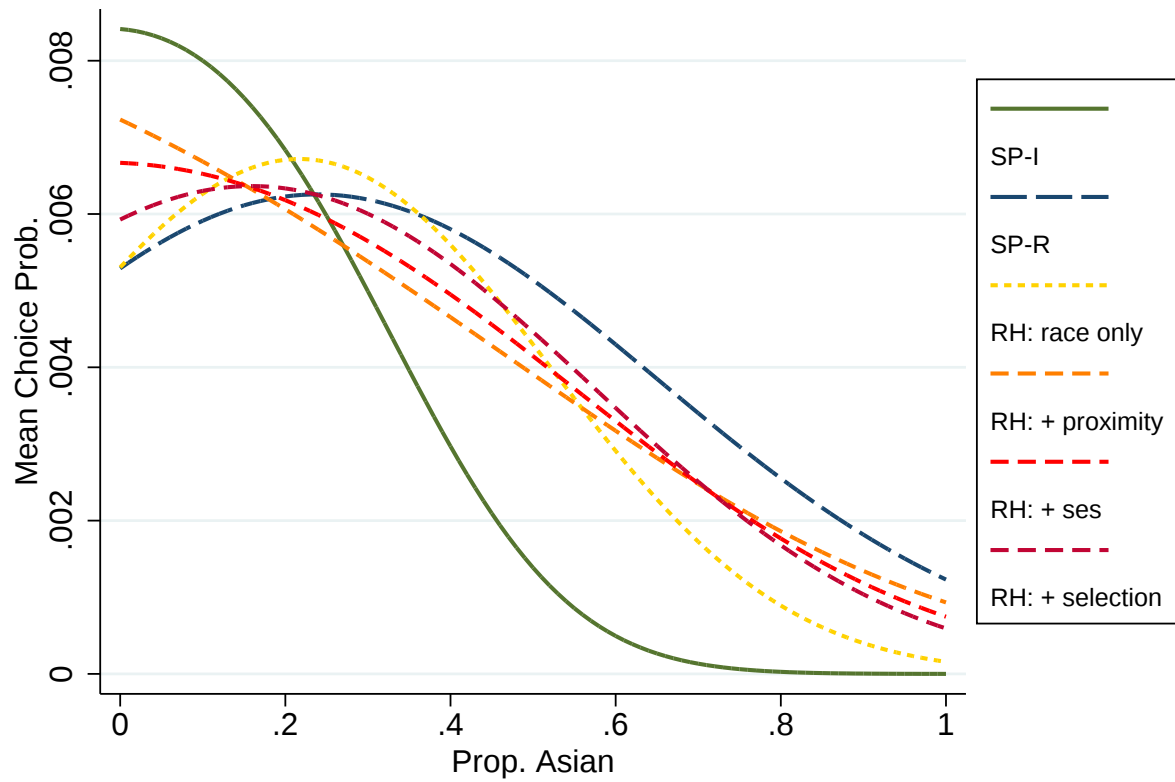


Figure 2.7: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino

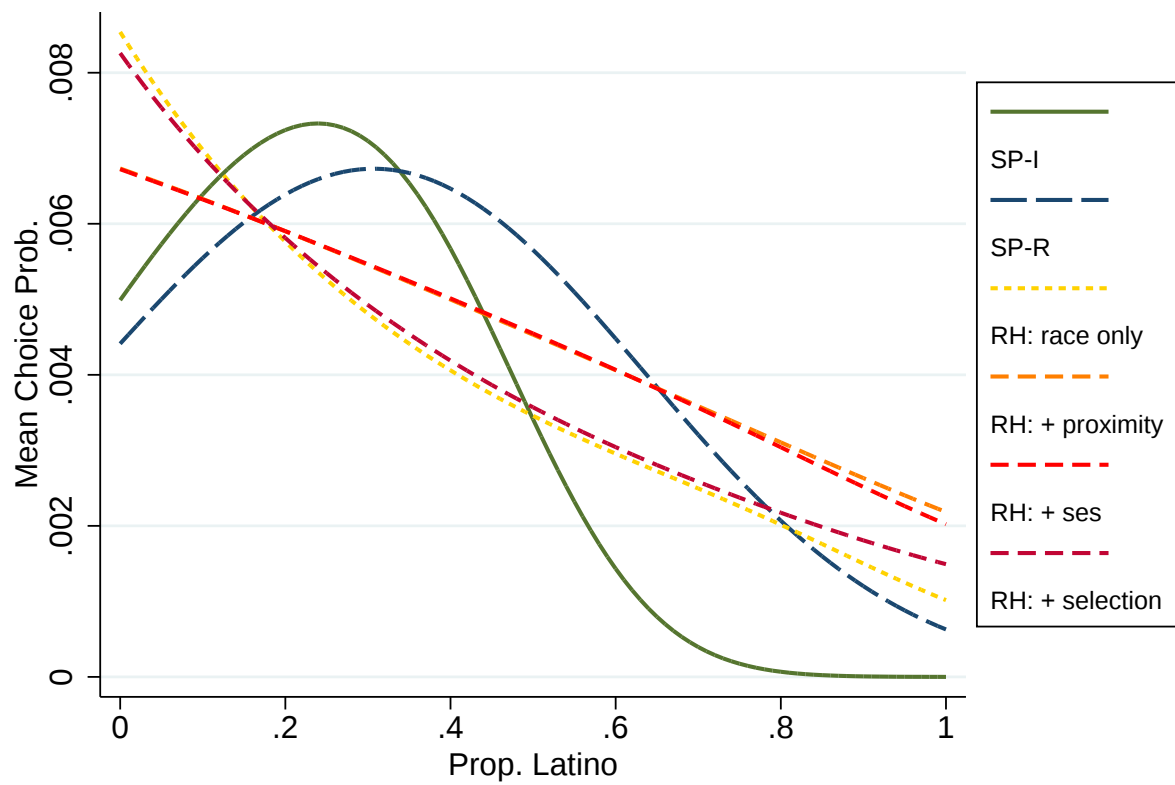


Figure 2.8: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black

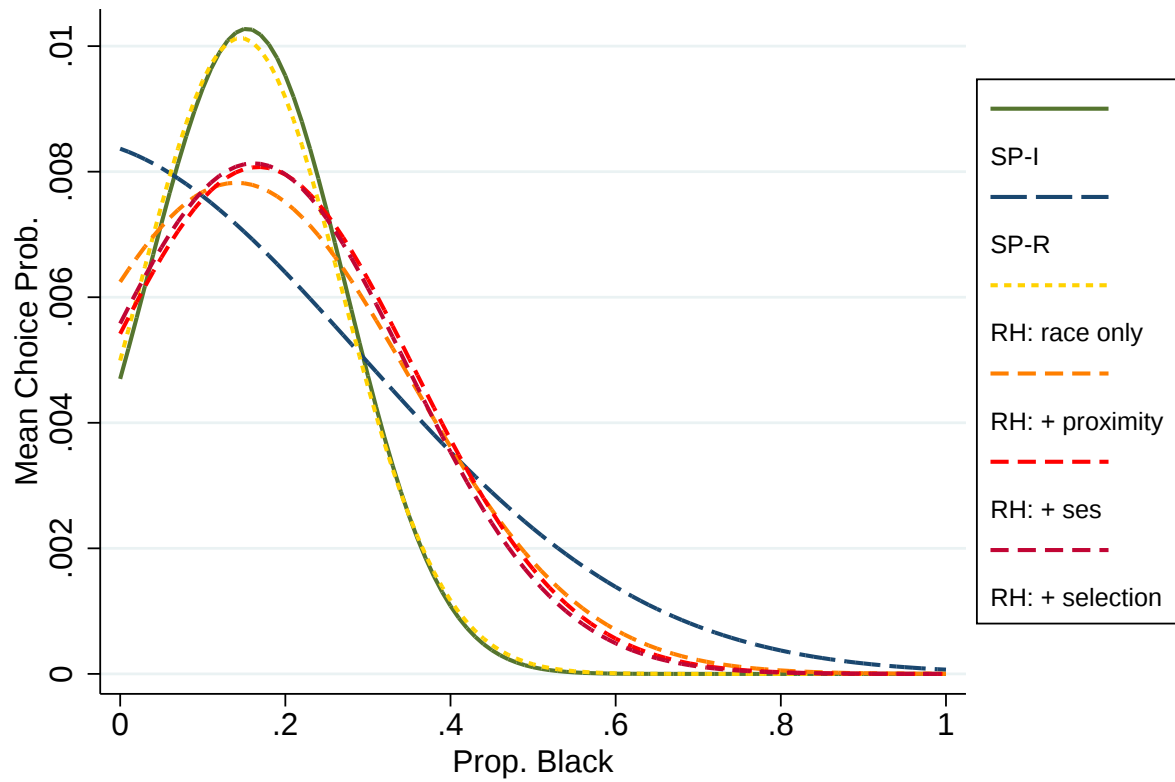


Figure 2.9: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White

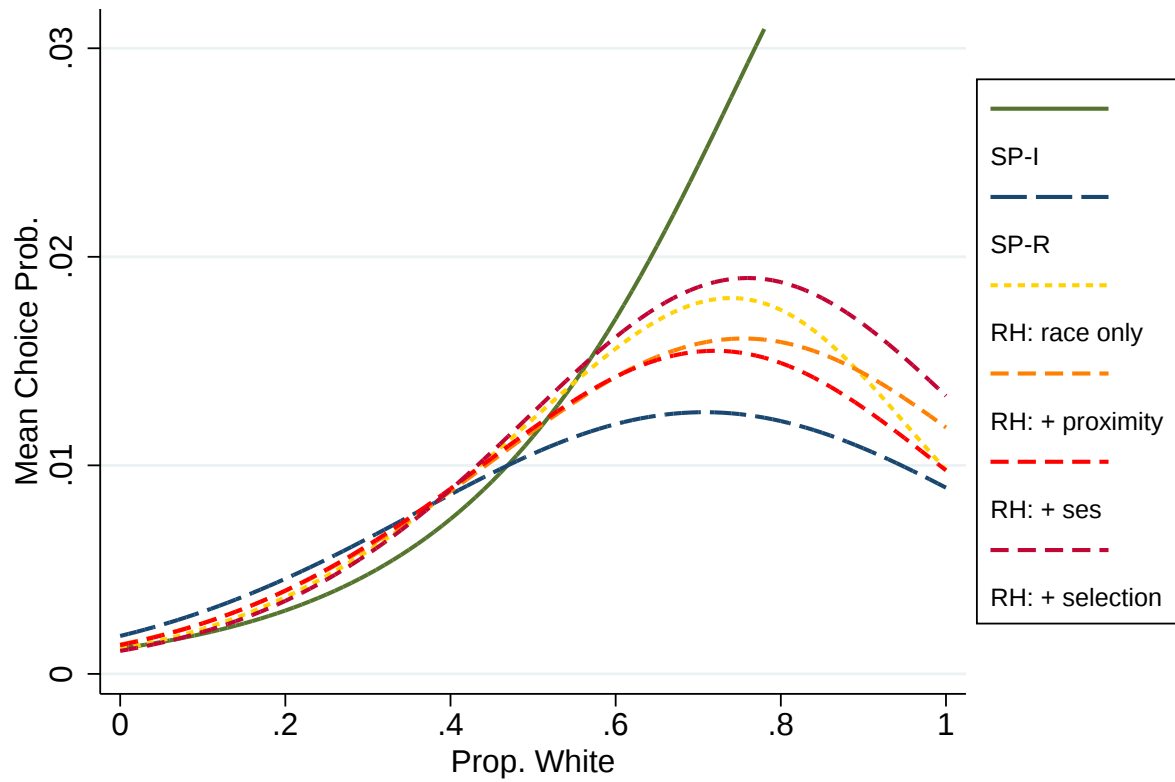


Figure 2.10: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian

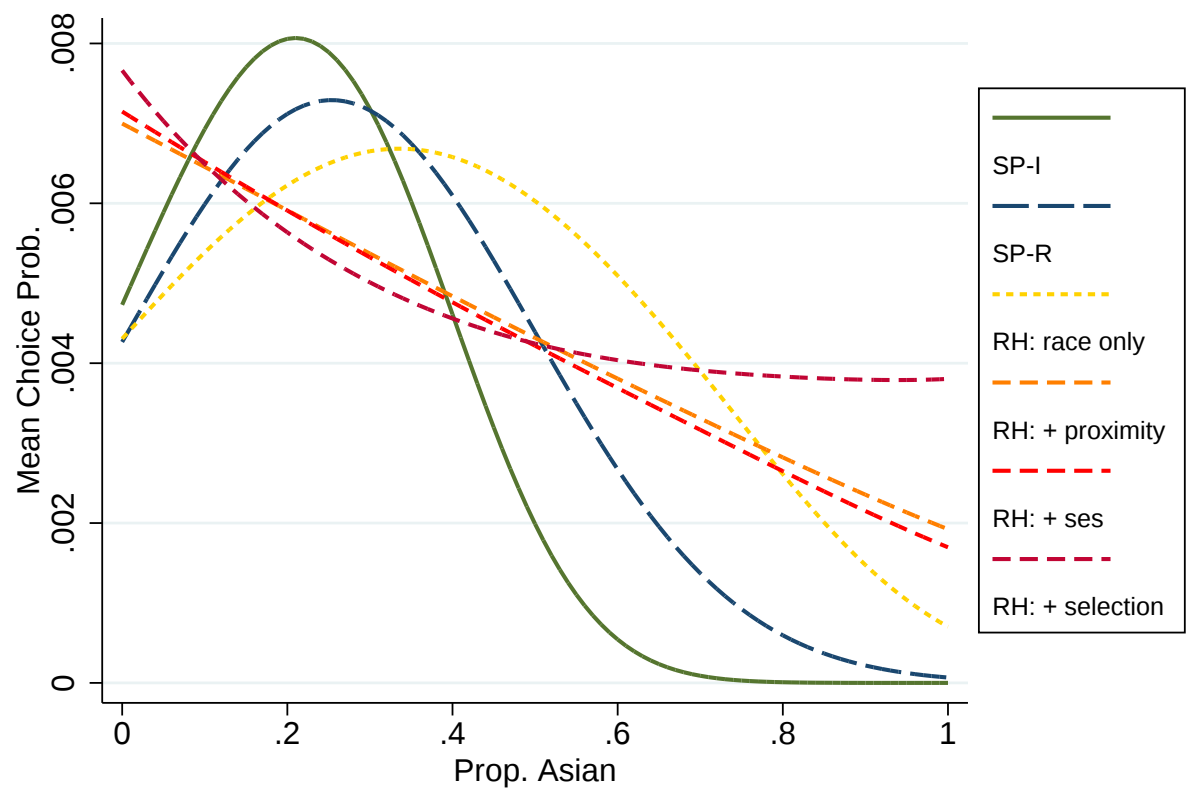


Figure 2.11: Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino

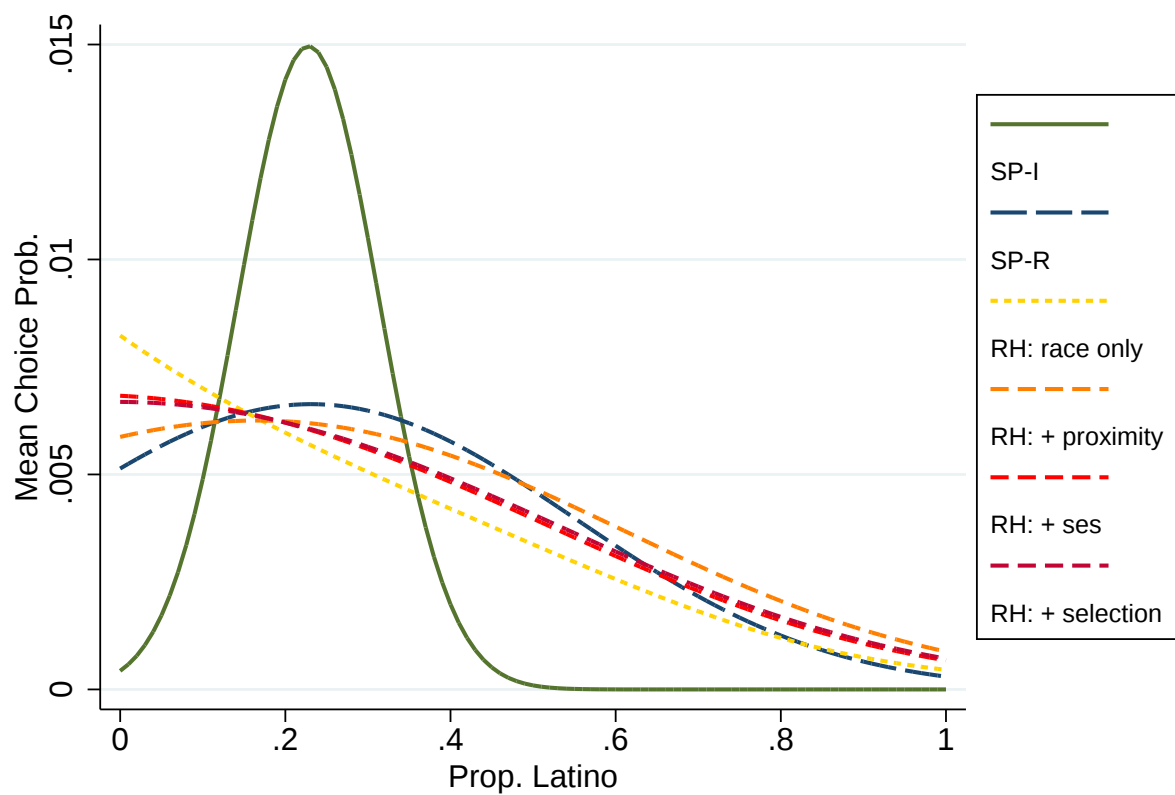


Figure 2.12: Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black

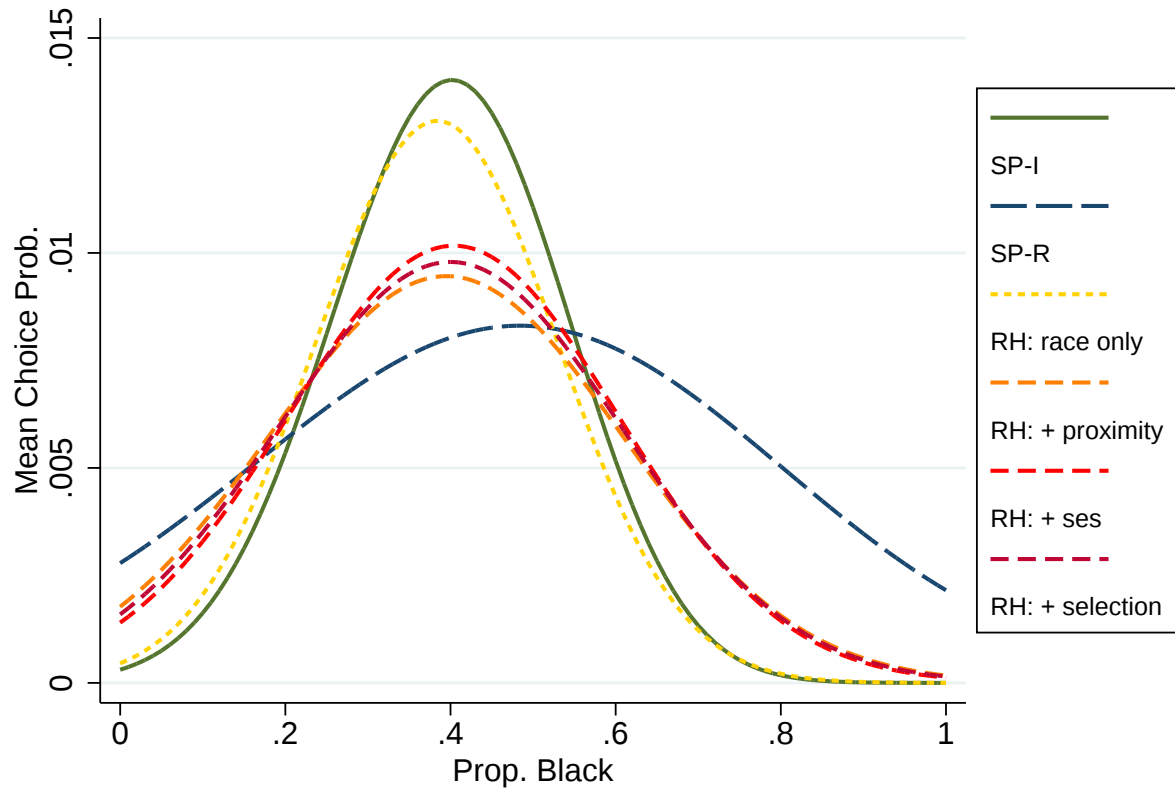


Figure 2.13: Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White

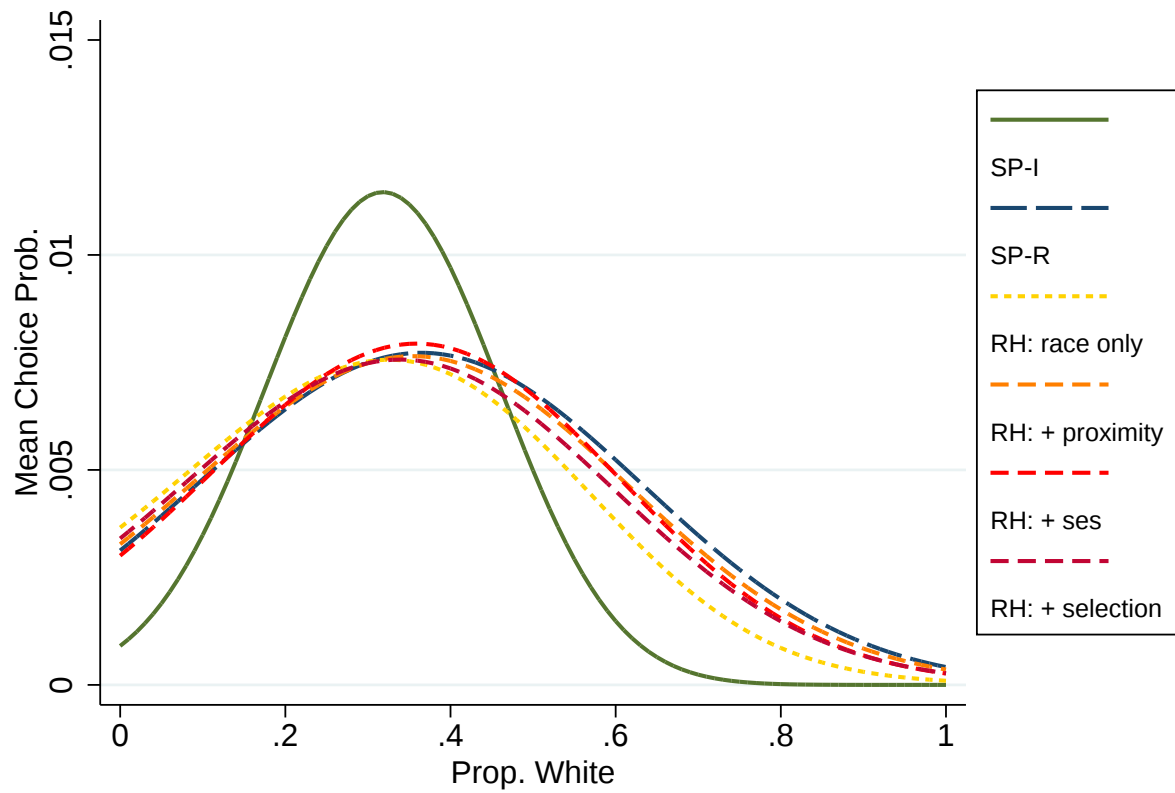


Figure 2.14: Black Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian

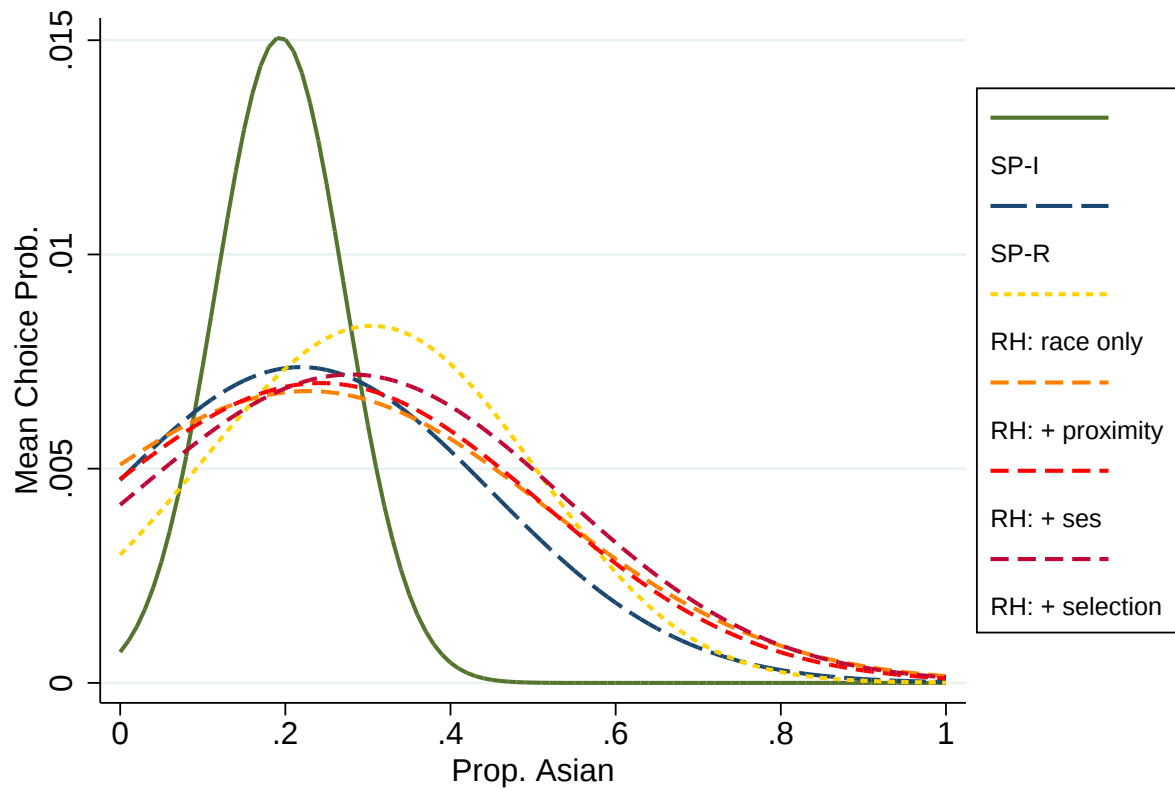


Figure 3.1: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian from Conditional Logit Models of Ranked SP Data

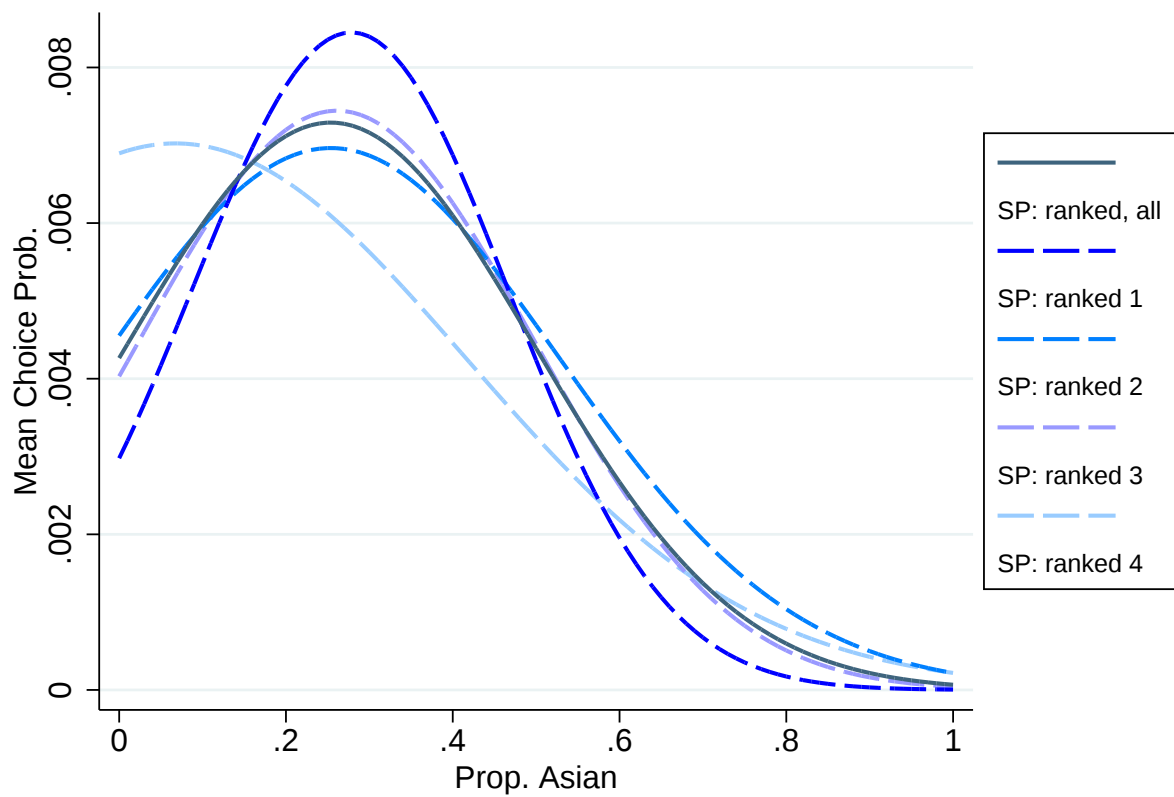


Figure 3.2: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black from Conditional Logit Models of Ranked SP Data

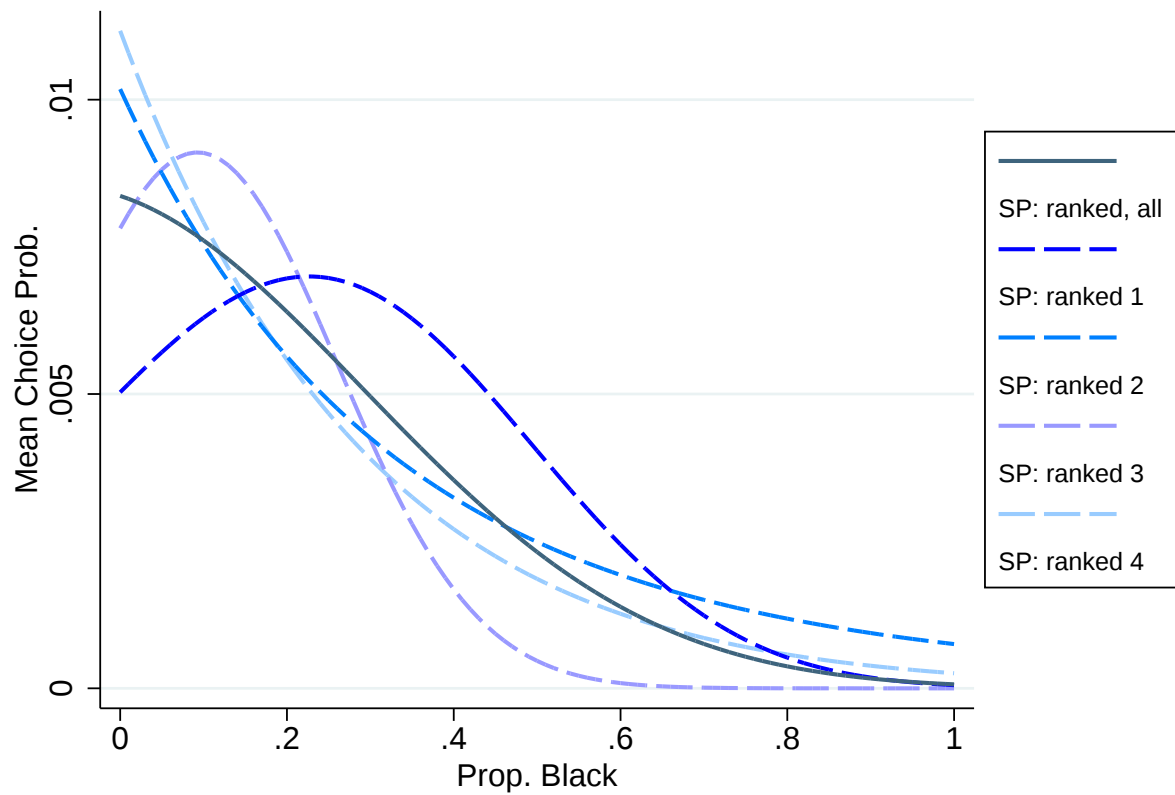


Figure 3.3: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino from Conditional Logit Models of Ranked SP Data

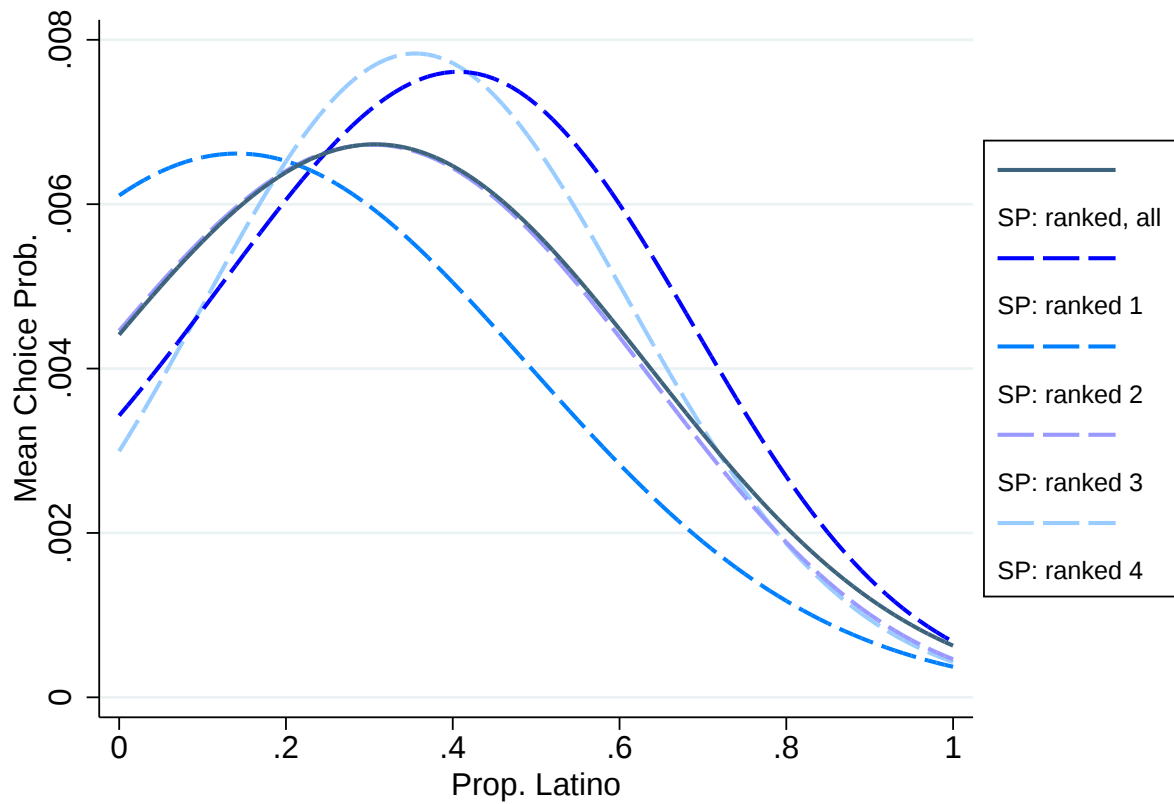


Figure 3.4: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White from Conditional Logit Models of Ranked SP Data

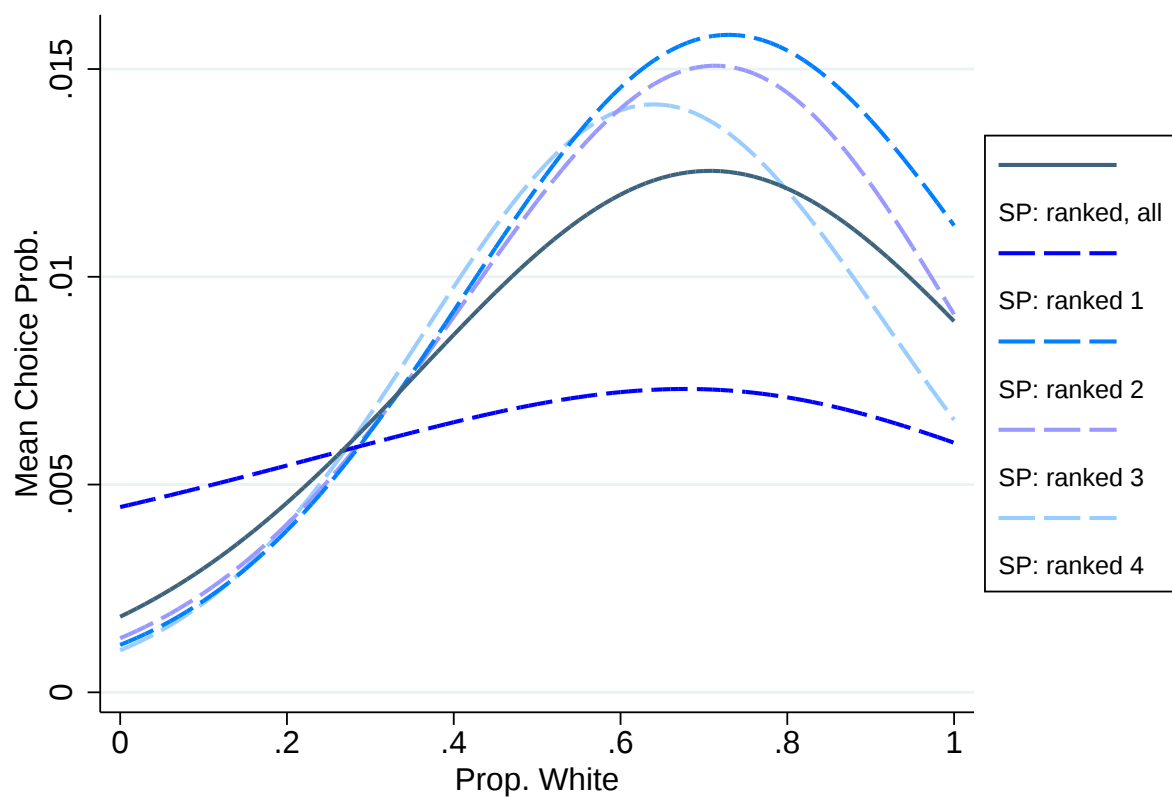


Figure 3.5: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Asian in Mixed Logit Models of Ranked SP Data

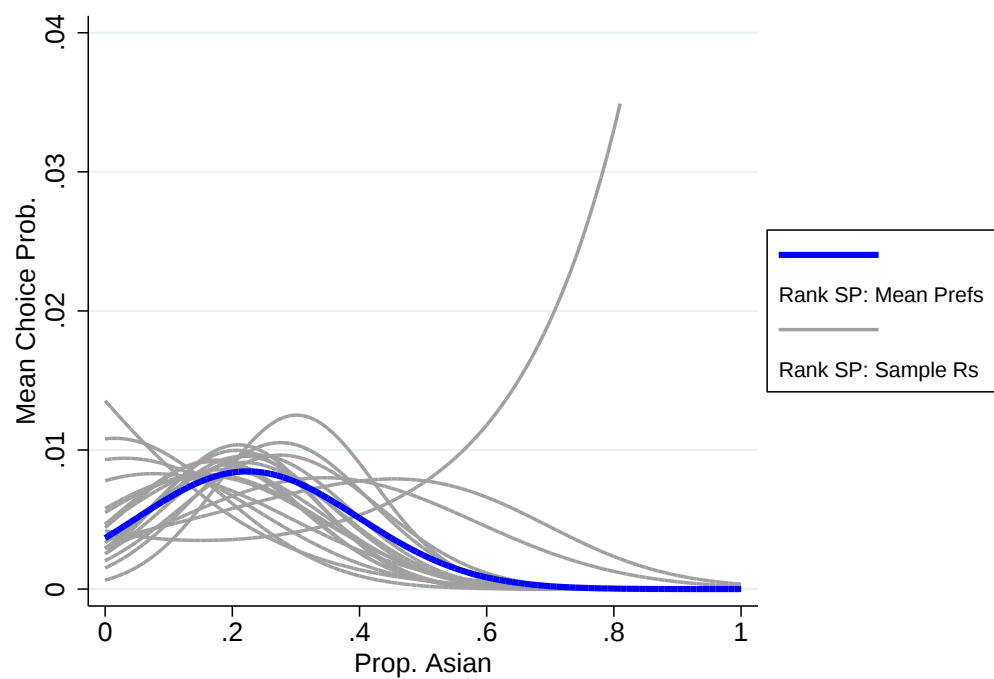


Figure 3.6: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Black, Mixed Logit Models of Ranked SP Data

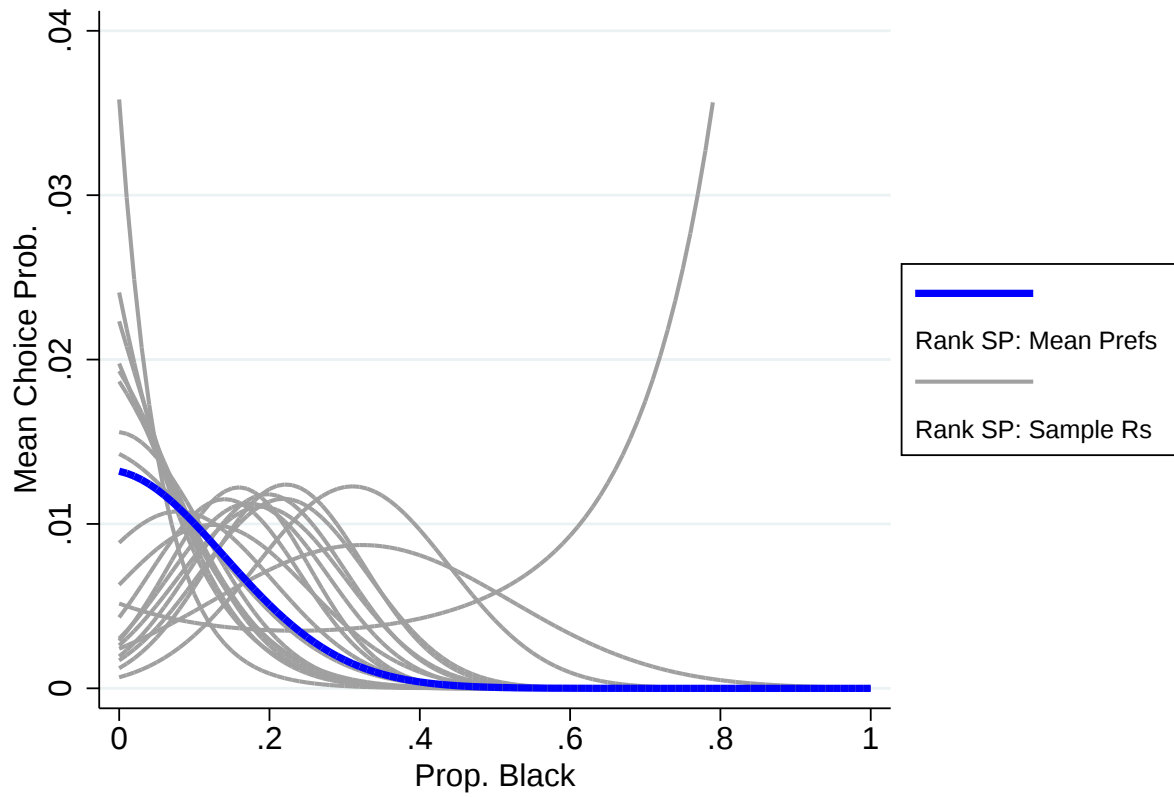


Figure 3.7: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion Latino, Mixed Logit Models of Ranked SP Data

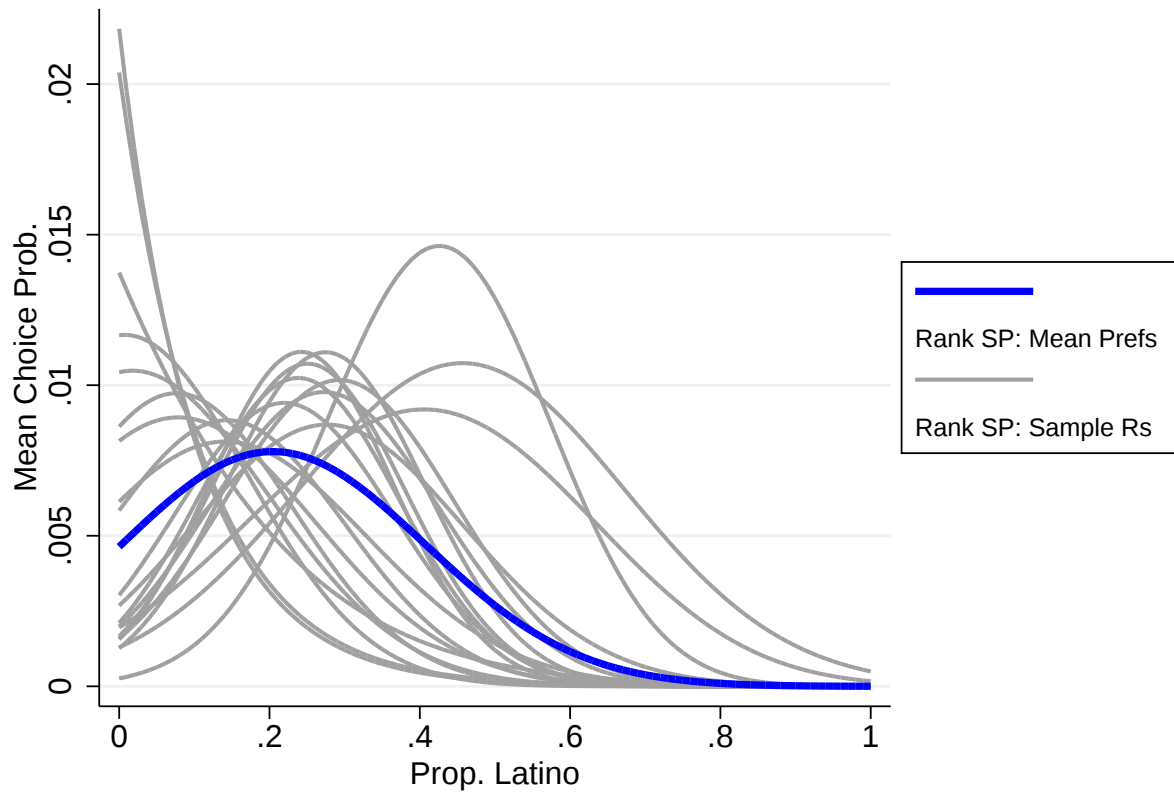


Figure 3.8: White Rs' Predicted Probability of Neighborhood Choice vs. Neighborhood Proportion White, Mixed Logit Models of Ranked SP Data

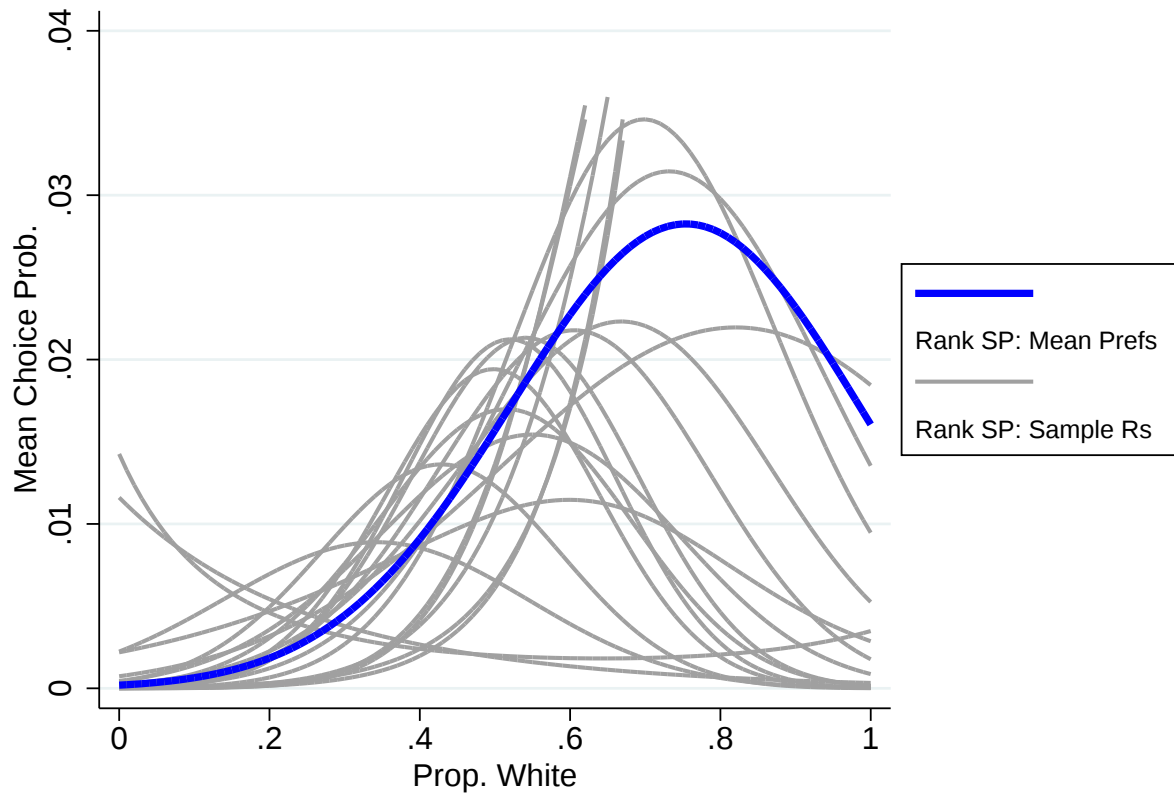


Figure 4.1: Typical Grid Set-Up for Schelling Models

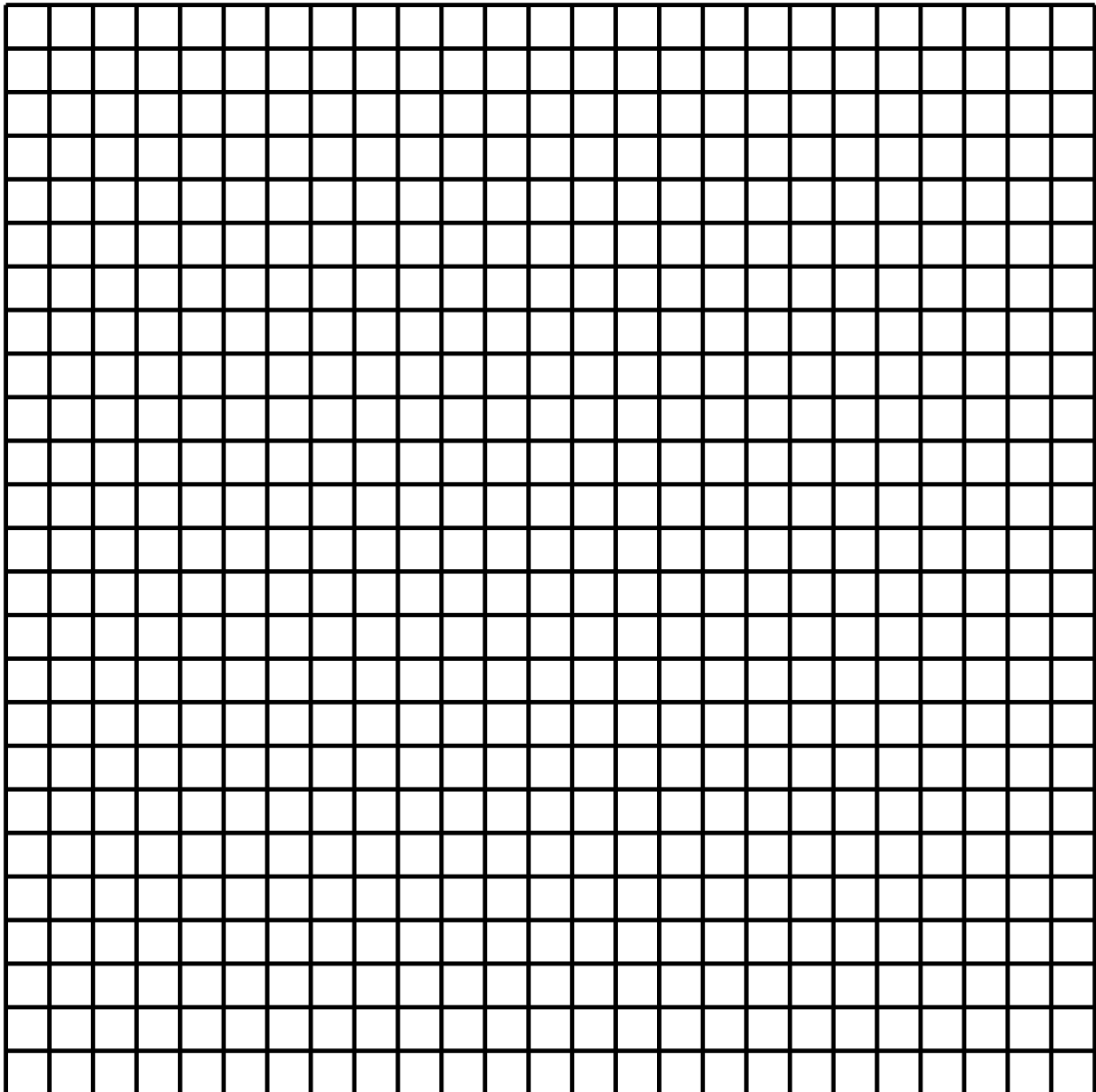
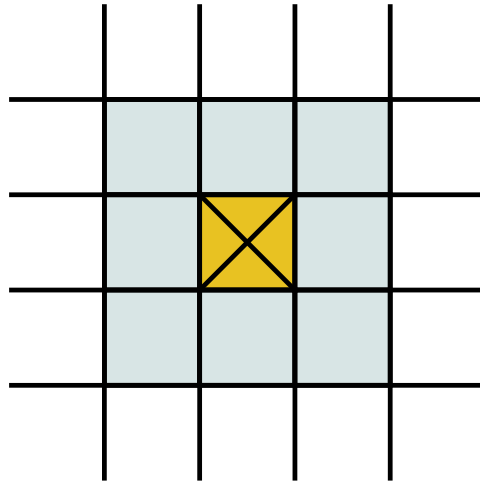
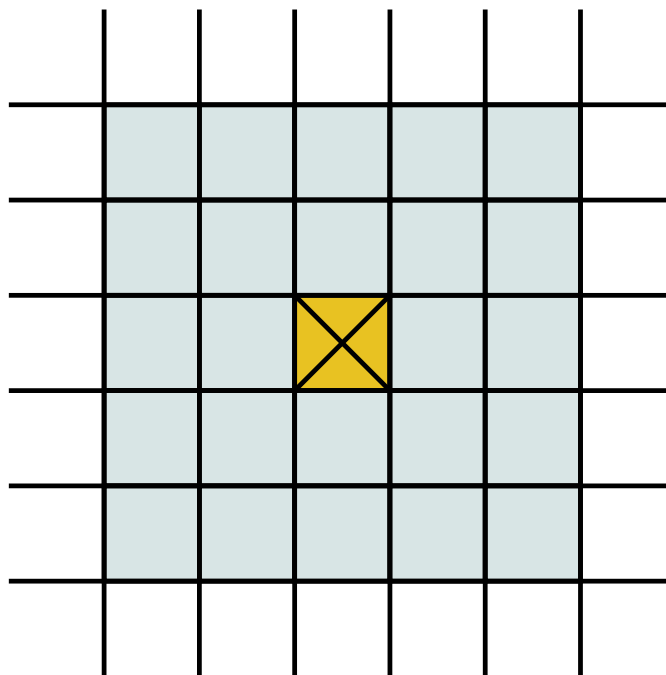


Figure 4.2: Neighborhood Definitions for Schelling Models



A) 3 x 3 Moore Neighborhood



B) 5 x 5 Moore Neighborhood

Figure 4.3: Initialization of City and Population Structure

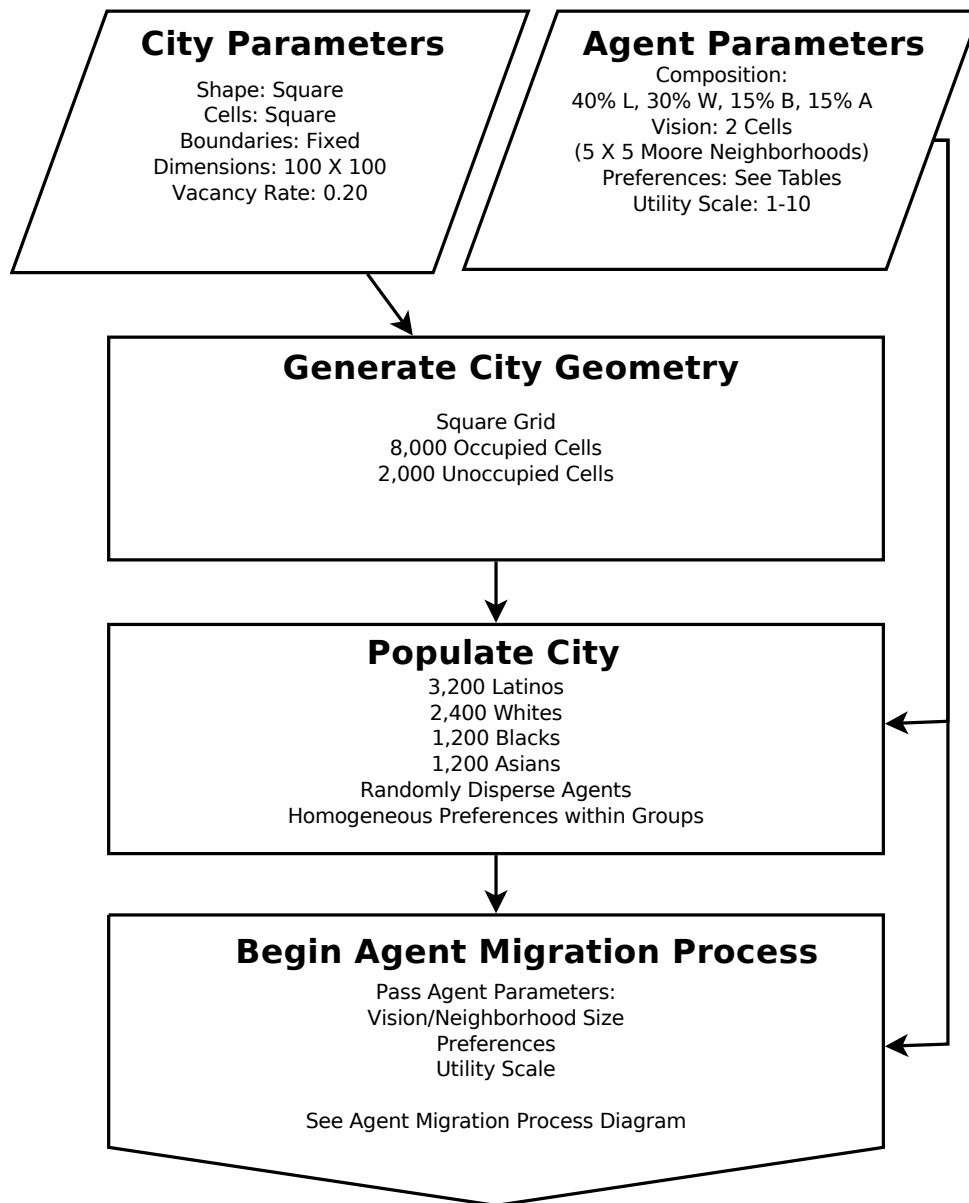


Figure 4.4: Agent Behavioral Specification

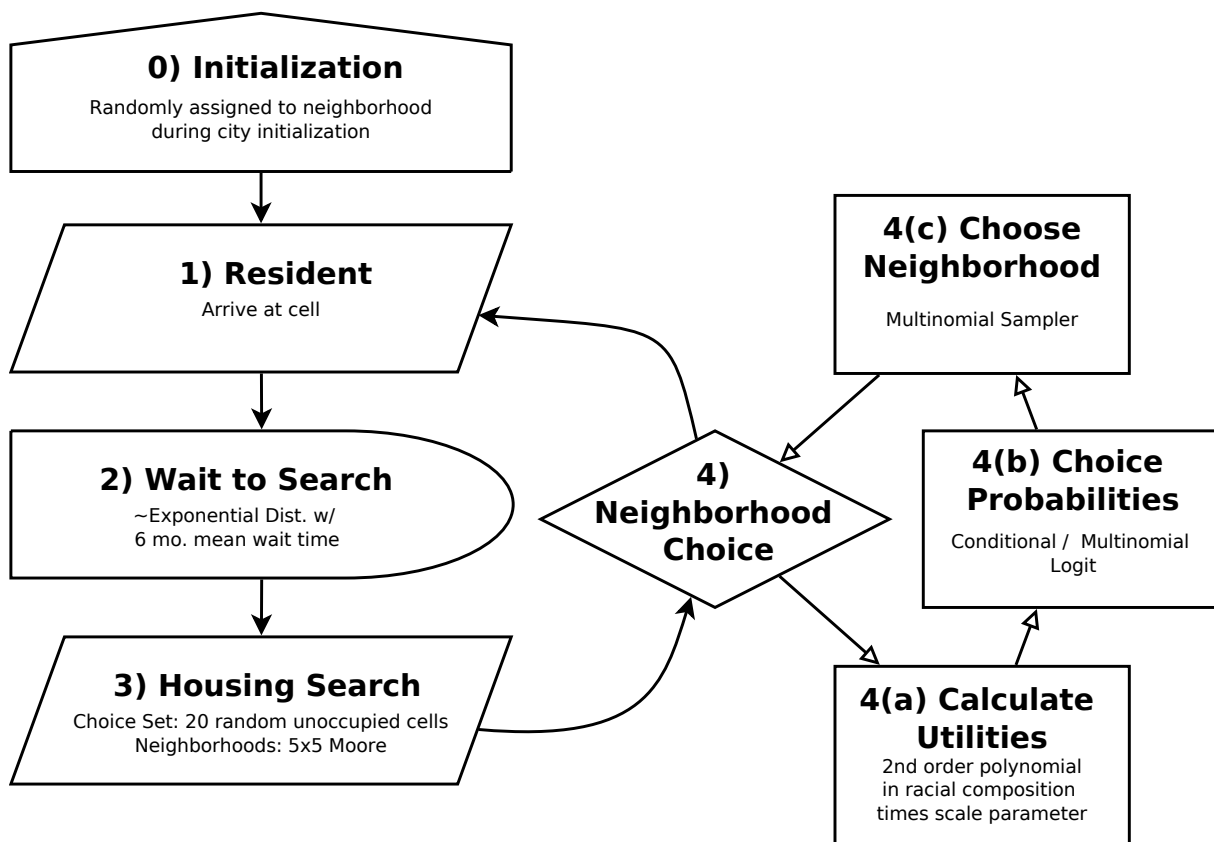


Figure 4.5: Theil's Entropy Index vs. Simulation Time, by Source of Racial Composition Coefficients and Scale Parameter

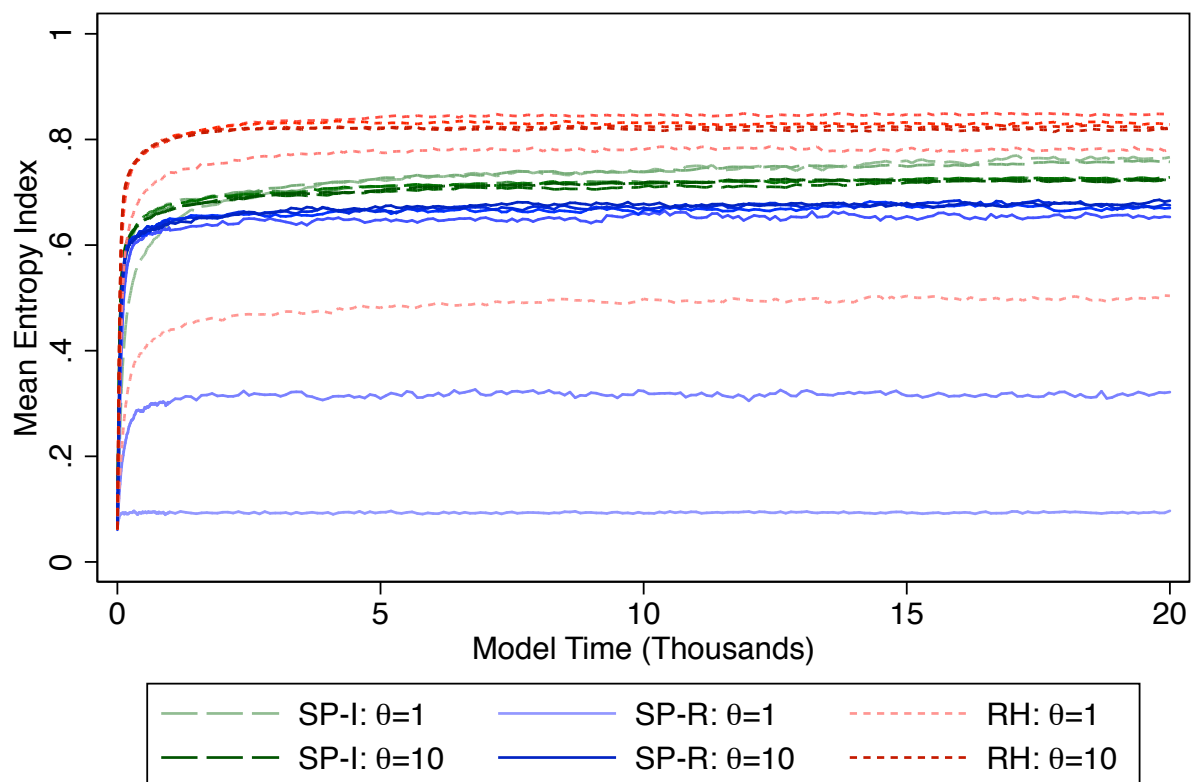


Figure 4.6: Dissimilarity Indices vs. Simulation Time, by Source of Racial Composition Coefficients and Scale Parameter

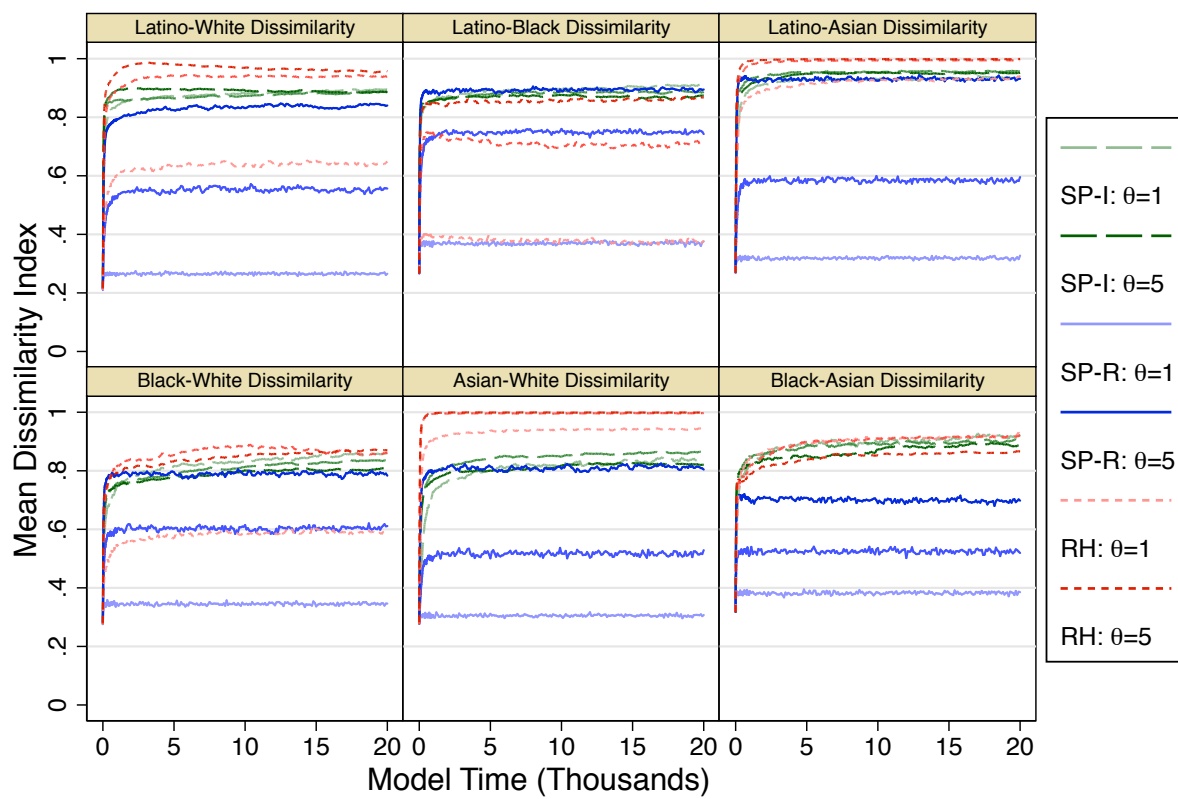
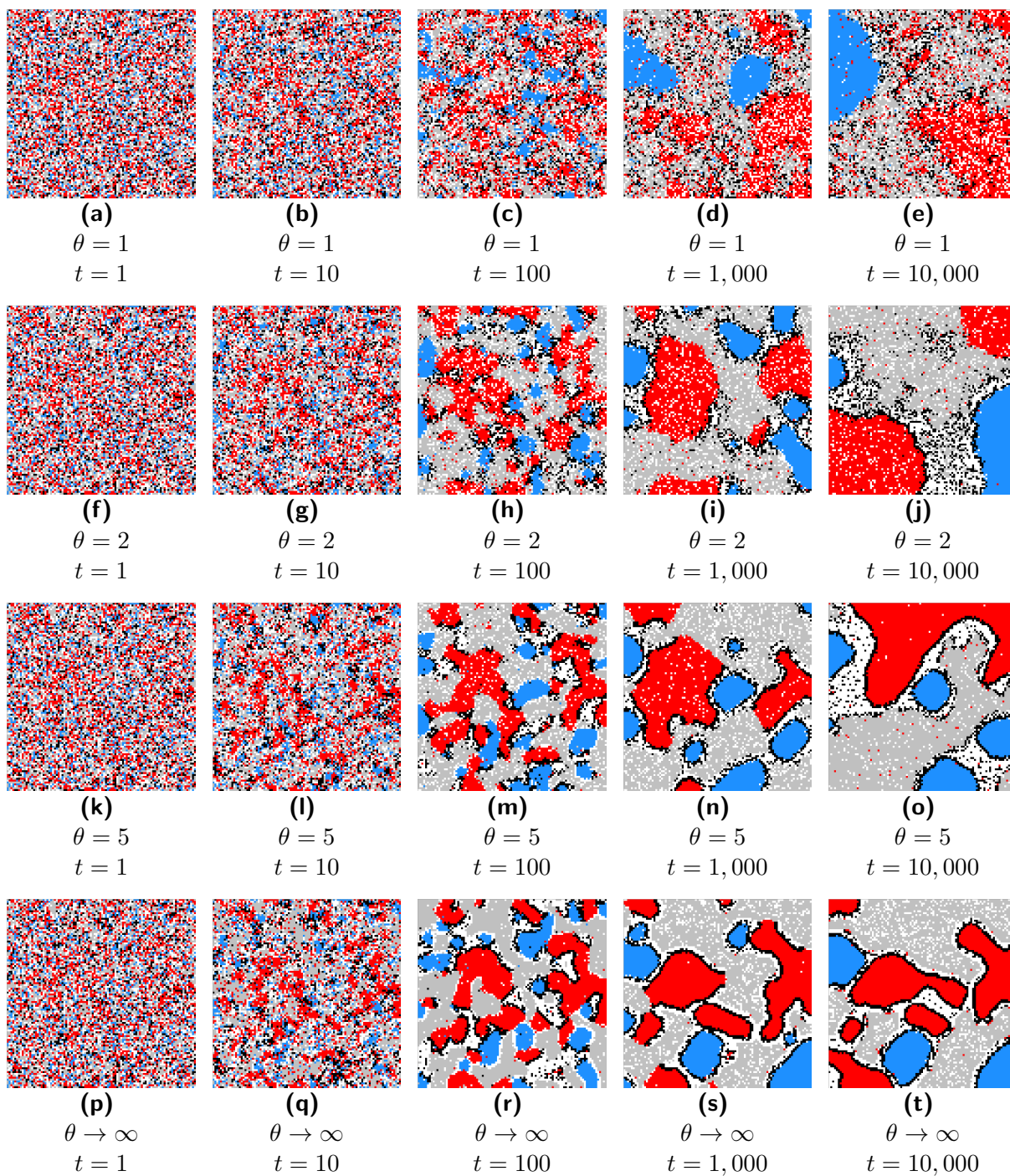
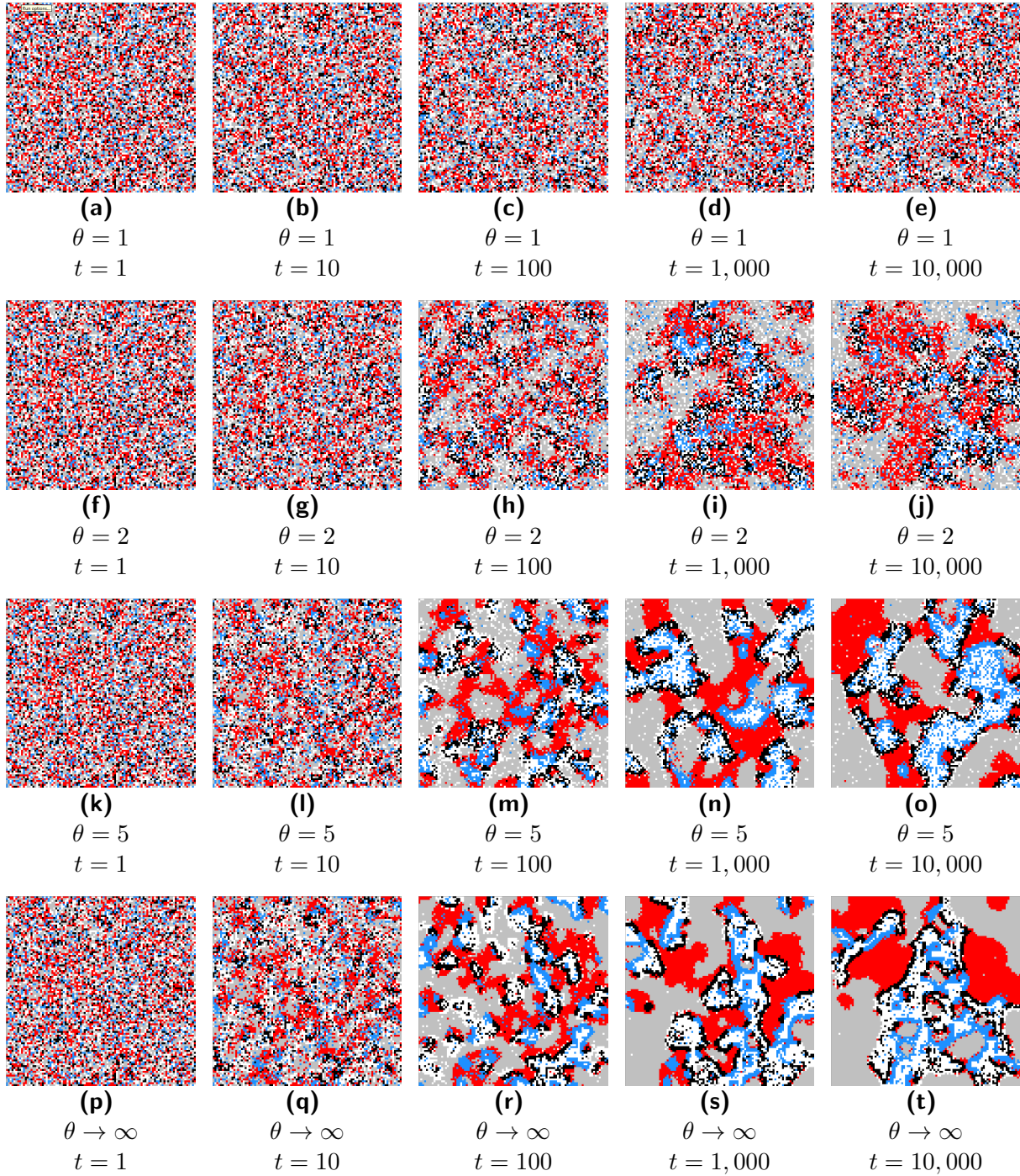


Figure 4.7: RH Coefficient Based Simulations



Silver = Latinos, Red = Whites, Black = Blacks, Blue = Asians

Figure 4.8: Rank SP Coefficient Based Simulations



Silver = Latinos, Red = Whites, Black = Blacks, Blue = Asians

Figure 4.9: Theil's Entropy Index vs. Scale Parameter (θ), by Source of Racial Composition Coefficients

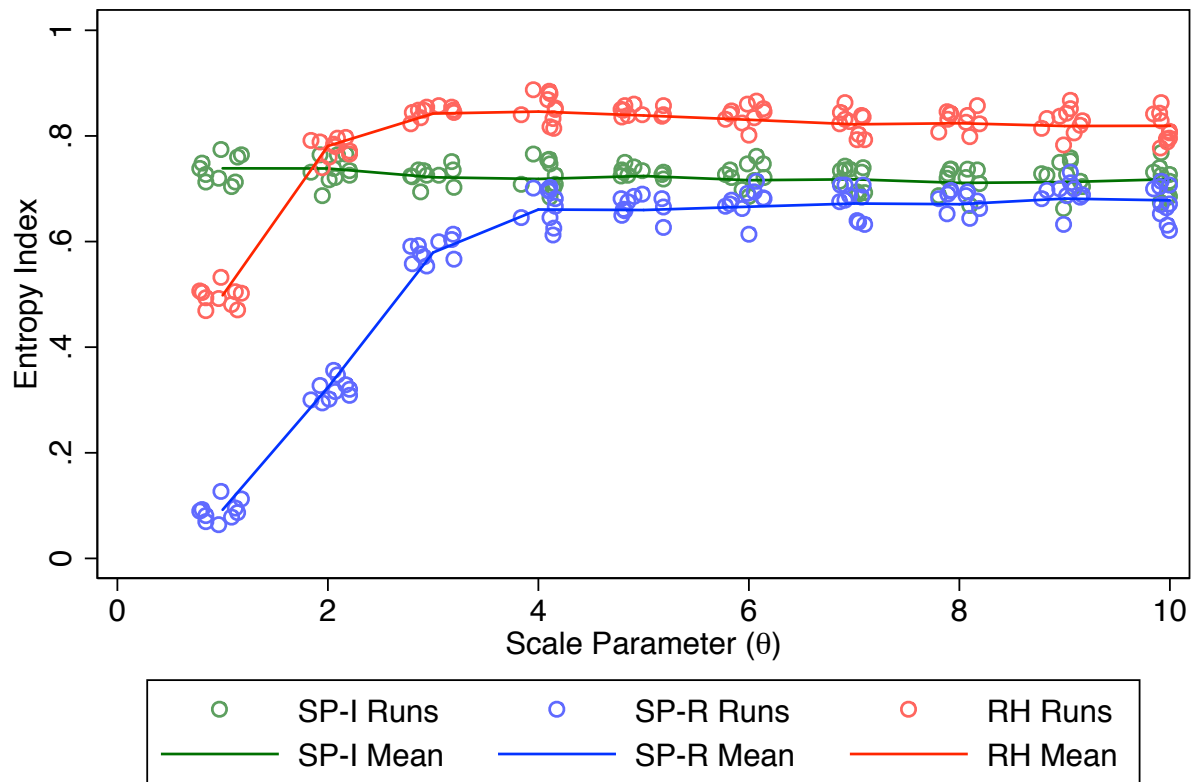
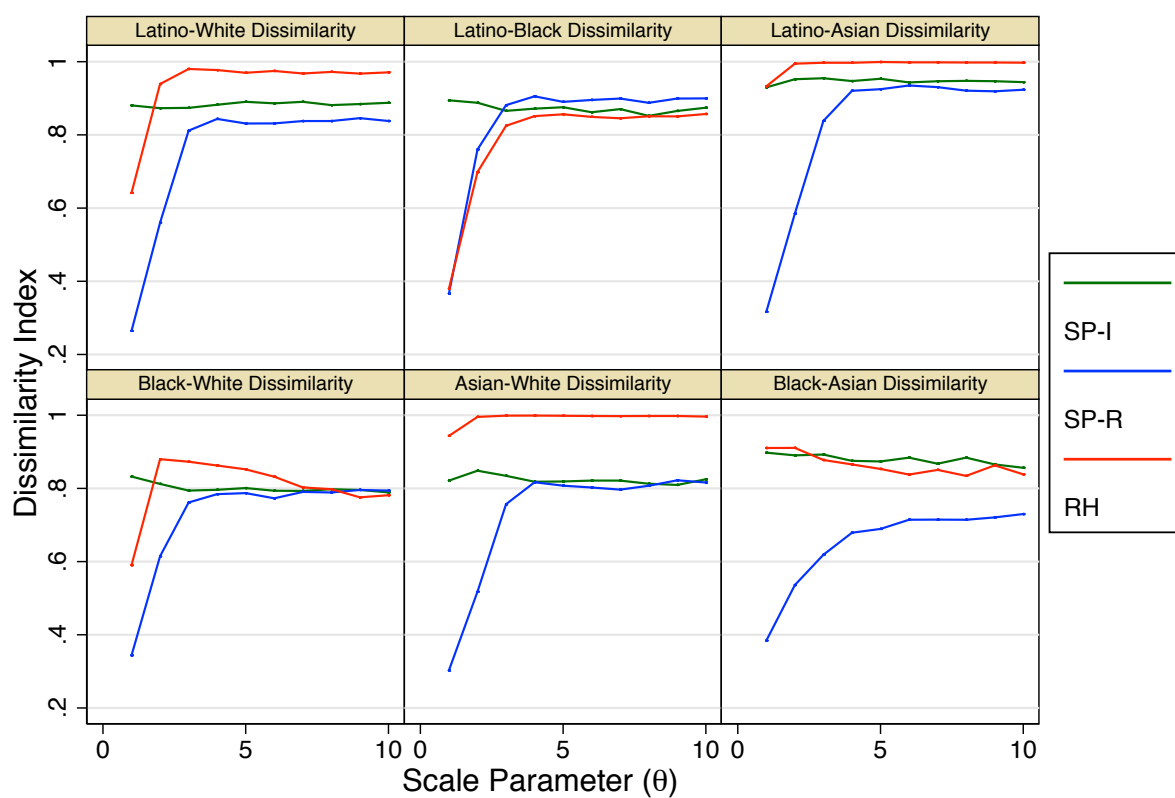


Figure 4.10: Dissimilarity Indices vs. Scale Parameter (θ), by Source of Racial Composition Coefficients



Bibliography

- Adelman, Robert 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: 209–228.
- Alba, Richard D. and John R. Logan. 1991. "Variations on Two Themes: Racial and Ethnic Patterns in the Attainment of Suburban Residence." *Demography* 28: 431–453.
- Alba, Richard D. and John R. Logan. 1993. "Minority Proximity to Whites in Suburbs: An Individual-Level Analysis of Segregation." *American Journal of Sociology* 98: 1388–1427.
- Allison, Paul D. and Nicholas A. Christakis. 1994. "Logit Models for Sets of Ranked Items." *Sociological Methodology* 24: 199–228.
- Axelrod, Robert M. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, N.J.: Princeton University Press.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115: 588–638.
- Bayer, Patrick and Robert McMillan. 2008. "Distinguishing Racial Preferences in the Housing Market: Theory and Evidence." In *Hedonic Methods in Housing Markets*, edited by Andrea Baranzini, José Ramirez, Caroline Schaerer, and Philippe Thalmann, pp. 225–244. New York: Springer.
- Bayer, Patrick, Robert McMillan, and Kim S. Rueben. 2004. "What Drives Racial Segregation? New Evidence Using Census Microdata." *Journal of Urban Economics* 56: 514–535.
- Ben-Akiva, Moshe, Mark Bradley, Takayuki Morikawa, Julian Benjamin, Thomas P. Novak, Harmen Oppewal, and V.M. Rao. 1994. "Combining revealed and stated preferences data." *Marketing Letters* 5: 335–349.
- Ben-Akiva, Moshe, Takayuki Morikawa, and Fumiaki Shiroishi. 1991. "Analysis of the reliability of preference ranking data." *Journal of Business Research* 23: 253–268.
- Ben-Akiva, Moshe E. and Steven R. Lerman. 1985. *Discrete Choice Analysis*. MIT Press.
- Benenson, Itzhak, Erez Hatna, and Ehud Or. 2009. "From Schelling to Spatially Explicit Modeling of Urban Ethnic and Economic Residential Dynamics." *Sociological Methods & Research* 37: 463–497.

- Bhat, Chandra R. 2001. "Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model." *Transportation Research Part B: Methodological* 35: 677–693.
- Bhat, Chandra R. 2003. "Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences." *Transportation Research Part B: Methodological* 37: 837–855.
- Bhat, Chandra R. and Saul Castelar. 2002. "A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area." *Transportation Research Part B: Methodological* 36: 593–616.
- Blalock, Hubert M. 1967. *Toward a Theory of Minority-Group Relations*. New York: Wiley.
- Blumer, Herbert. 1958. "Race Prejudice as a Sense of Group Position." *The Pacific Sociological Review* 1: 3–7.
- Bobo, Lawrence and Camille L. Zubrinski. 1996. "Attitudes on Residential Integration: Perceived Status Differences, Mere In-Group Preference, or Racial Prejudice?" *Social Forces* 74: 883–909.
- Bobo, Lawrence D. 1999. "Prejudice as Group Position: Microfoundations of a Sociological Approach to Racism and Race Relations." *Journal of Social Issues* 55: 445–472.
- Bobo, Lawrence D. and Devon Johnson. 2000. "Racial Attitudes in a Prismatic Metropolis: Mapping Identity, Stereotypes, Competition, and Views on Affirmative Action." In *Prismatic Metropolis: Inequality in Los Angeles*, edited by Lawrence D. Bobo, Melvin L. Oliver, James H. Johnson Jr., and Abel Valenzuela, volume 5 of *Multi-City of Urban Inequality*. New York: Russell Sage Foundation.
- Boustan, Leah Platt. 2010. "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration." *The Quarterly Journal of Economics* 125: 417–443.
- Boyd, Monica. 1989. "Family and Personal Networks in International Migration: Recent Developments and New Agendas." *International Migration Review* 23: 638.
- Brownstone, David, David S. Bunch, and Kenneth Train. 2000. "Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles." *Transportation Research Part B: Methodological* 34: 315–338.
- Bruch, Elizabeth E. 2012. "Understanding the Link between Neighborhood Desires and Neighborhood Attainment." In *Population Association of America Annual Meeting*, San Francisco, CA. Population Association of America. Session 134: Residential Mobility, Suburbanization, and Exurbanization.
- Bruch, Elizabeth E. 2014. "How Population Structure Shapes Neighborhood Segregation." *American Journal of Sociology* 119: 1221–1278.

- Bruch, Elizabeth E. and Robert D. Mare. 2006. "Neighborhood Choice and Neighborhood Change." *American Journal of Sociology* 112: 667–709.
- Bruch, Elizabeth E. and Robert D. Mare. 2009. "Preferences and Pathways to Segregation: Reply to Van De Rijt, Siegel, and Macy." *American Journal of Sociology* 114: 1181–1198.
- Bruch, Elizabeth E. and Robert D. Mare. 2012. "Methodological Issues in the Analysis of Residential Preferences, Residential Mobility, and Neighborhood Change." *Sociological Methodology* 42: 103–154.
- Chang, Jae Bong and Jayson L. Lusk. 2011. "Mixed Logit Models: Accuracy and Software Choice." *Journal of Applied Econometrics* 26: 167–172.
- Charles, Camille Zubrinsky. 2000. "Neighborhood Racial-Composition Preferences: Evidence from a Multiethnic Metropolis." *Social Problems* 47: 379–407.
- Charles, Camille Zubrinsky. 2006. *Won't You Be My Neighbor?: Race, Class, and Residence in Los Angeles*. Russell Sage Foundation Publications.
- Charles, Camille Zubrinsky. 2007. "Comfort Zones-immigration, Acculturation, and the Neighborhood Racial-Composition Preferences of Latinos and Asians." *Du Bois Review: Social Science Research on Race* 4: 41–77.
- Chintagunta, Pradeep K., Dipak C. Jain, and Naufel J. Vilcassim. 1991. "Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data." *Journal of Marketing Research* 28: 417–428.
- Choldin, Harvey M. 1973. "Kinship Networks in the Migration Process." *International Migration Review* 7: 163.
- Clark, William A.V. 2002. "Ethnic Preferences and Ethnic Perceptions in Multi-Ethnic Settings." *Urban Geography* 23: 237–256.
- Clark, William A. V. 1991. "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model." *Demography* 28: 1–19.
- Clark, William A. V. 1992. "Residential Preferences and Residential Choices in a Multiethnic Context." *Demography* 29: 451–466.
- Clark, William A. V. 2009. "Changing Residential Preferences across Income, Education, and Age: Findings from the Multi-City Study of Urban Inequality." *Urban Affairs Review* 44: 334–355.
- Clark, William A. V. and Mark Fossett. 2008. "Understanding the social context of the Schelling segregation model." *Proceedings of the National Academy of Sciences* 105: 4109–4114.
- Coleman, James S. 1975. "Trends in School Segregation, 1968-73." Technical report, The Urban Institute, Washington, D.C.

- Collins, Chiquita A. and David R. Williams. 1999. "Segregation and Mortality: The Deadly Effects of Racism?" *Sociological Forum* 14: 495–523.
- Crowder, Kyle, Jeremy Pais, and Scott J South. 2012. "Neighborhood Diversity, Metropolitan Constraints, and Household Migration." *American Sociological Review* 77: 325–353.
- Crowder, Kyle and Scott J. South. 2005. "Race, Class, and Changing Patterns of Migration between Poor and Nonpoor Neighborhoods." *American Journal of Sociology* 110: 1715–1763.
- Crowder, Kyle D. 2000. "The Racial Context of White Mobility: An Individual-Level Assessment of the White Flight Hypothesis." *Social Science Research* 29: 223–257. 2.
- Crowder, Kyle D. 2001. "Racial Stratification in the Actuation of Mobility Expectations: Microlevel Impacts of Racially Restrictive Housing Markets." *Social Forces* 79: 1377–1396.
- Crowder, Kyle D. and Scott J. South. 2008. "Spatial Dynamics of White Flight: The Effects of Local and Extralocal Racial Conditions on Neighborhood Out-Migration." *American Sociological Review* 73: 792–812.
- Crowder, Kyle D., Scott J. South, and Erick Chavez. 2006. "Wealth, Race, and Inter-Neighborhood Migration." *American Sociological Review* 71: 72–94.
- de Palma, André, Nathalie Picard, and Paul Waddell. 2007. "Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region." *Journal of Urban Economics* 62: 204–230.
- Denton, Nancy A and Douglas S Massey. 1988. "Residential Segregation of Blacks, Hispanics, and Asians by Socioeconomic Status and Generation." *Social Science Quarterly* 69: 797–817.
- Ellen, Ingrid Gould. 2000. *Sharing America's Neighborhoods*. Cambridge, MA: Harvard University Press.
- Ellis, Mark, Steven R. Holloway, Richard Wright, and Christopher S. Fowler. 2012. "Agents of Change: Mixed-Race Households and the Dynamics of Neighborhood Segregation in the United States." *Annals of the Association of American Geographers* 102: 549–570.
- Emerson, Michael O., Karen J. Chai, and George Yancey. 2001. "Does Race Matter in Residential Segregation? Exploring the Preferences of White Americans." *American Sociological Review* 66: 922–935.
- Epstein, Joshua M. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton: Princeton University Press.
- Epstein, Joshua M., Robert Axtell, and 2050 Project. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, D.C.: Brookings Institution Press.

- Farley, Reynolds, Suzanne Bianchi, and Diane Colasanto. 1979. "Barriers to the Racial Integration of Neighborhoods: The Detroit Case." *The Annals of the American Academy of Political and Social Science* 441: 97–113.
- Farley, Reynolds, Elaine L. Fielding, and Maria Krysan. 1997. "The Residential Preferences of Blacks and Whites: A Four-Metropolis Analysis." *Housing Policy Debate* 8: 763–800.
- Farley, Reynolds, Howard Schuman, Suzanne Bianchi, Diane Colasanto, and Shirley Hatchett. 1978. "'Chocolate City, Vanilla Suburbs': Will the Trend Toward Racially Separate Communities Continue?" *Social Science Research* 7: 319–344.
- Farley, Reynolds, C. Steeh, T. Jackson, Maria Krysan, and K. Reeves. 1993. "Continued Racial Residential Segregation in Detroit: 'Chocolate City, Vanilla Suburbs' Revisited." *Journal of Housing Research* 4: 1–38.
- Farley, Reynolds, Charlotte Steeh, Maria Krysan, Tara Jackson, and Keith Reeves. 1994. "Stereotypes and Segregation: Neighborhoods in the Detroit Area." *The American Journal of Sociology* 100: 750–780.
- Feagin, Joe R. and Melvin P. Sikes. 1994. *Living with Racism: The Black Middle-Class Experience*. Beacon Press.
- Fossett, Mark. 2006a. "Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations Using Simulation Analysis." *The Journal of Mathematical Sociology* 30: 185–273.
- Fossett, Mark. 2006b. "Including Preference and Social Distance Dynamics in Multi-Factor Theories of Segregation." *The Journal of Mathematical Sociology* 30: 289–298.
- Fossett, Mark. 2011. "Generative Models of Segregation: Investigating Model-Generated Patterns of Residential Segregation by Ethnicity and Socioeconomic Status." *The Journal of Mathematical Sociology* 35: 114–145.
- Fossett, Mark and David R. Dietrich. 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 and Planning B: Planning and Design* 36: 149–169.
- Fossett, Mark and Warren Warren. 2005. "Overlooked Implications of Ethnic Preferences for Residential Segregation in Agent-based Models." *Urban Studies* 42: 1893–1917.
- Freeman, Lance. 2000. "Minority Housing Segregation: A Test of Three Perspectives." *Journal of Urban Affairs* 22: 15–35.
- Frey, William H. 1979. "Central City White Flight: Racial and Nonracial Causes." *American Sociological Review* 44: 425–448.
- Galster, George C. 1987. "The Ecology of Racial Discrimination in Housing An Exploratory Model." *Urban Affairs Review* 23: 84–107.

- Galster, George C. 1988. "Residential segregation in American cities: A contrary review." *Population Research and Policy Review* 7: 93–112.
- Galster, George C. and W. Mark Keeney. 1988. "Race, Residence, Discrimination, and Economic Opportunity Modeling the Nexus of Urban Racial Phenomena." *Urban Affairs Review* 24: 87–117.
- Geolytics. 2003. *Neighborhood Change Database*. East Brunswick, NJ: Geolytics, Inc.
- Greene, William H. and David A. Hensher. 2003. "A latent class model for discrete choice analysis: contrasts with mixed logit." *Transportation Research Part B: Methodological* 37: 681–698.
- Greenwood, Michael J. 1985. "Human Migration: Theory, Models, and Empirical Studies." *Journal of Regional Science* 25: 521–544.
- Greenwood, Michael J. 1997. "Chapter 12 Internal Migration in Developed Countries." In *Handbook of Population and Family Economics*, volume 1B, pp. 647–720. Elsevier.
- Harding, David J. 2003. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy." *American Journal of Sociology* 109: 676–719.
- Harris, David R. 1999. "'Property Values Drop When Blacks Move in, Because...': Racial and Socioeconomic Determinants of Neighborhood Desirability." *American Sociological Review* 64: 461–479.
- Harris, David R. 2001. "Why Are Whites and Blacks Averse to Black Neighbors?" *Social Science Research* 30: 100–116.
- Heckman, James J. 1981. "Heterogeneity and State Dependence." In *Studies in Labor Markets*, edited by Sherwin Rosen, pp. 91–140. Chicago: University of Chicago Press.
- Heckman, James J. 1998. "Detecting Discrimination." *The Journal of Economic Perspectives* 12: 101–116.
- Hensher, David A. and Mark Bradley. 1993. "Using stated response choice data to enrich revealed preference discrete choice models." *Marketing Letters* 4: 139–151.
- Hensher, David A. and William H. Greene. 2003. "The Mixed Logit Model: The State of Practice." *Transportation* 30: 133–176.
- Hensher, David A., Jordan J. Louviere, and Joffre Swait. 1998. "Combining sources of preference data." *Journal of Econometrics* 89: 197–221.
- Hensher, David A., John M. Rose, and William H. Greene. 2008. "Combining RP and SP data: biases in using the nested logit 'trick' – contrasts with flexible mixed logit incorporating panel and scale effects." *Journal of Transport Geography* 16: 126–133.

- Hole, A. R. 2007. "Fitting mixed logit models by using maximum simulated likelihood." *Stata Journal* 7: 388–401.
- Hole, Arne Risa. 2011. "A comment on 'Mixed logit models: Accuracy and software choice.'." Working Paper 2011011, The University of Sheffield, Department of Economics.
- Holloway, Steven R., Mark Ellis, Richard Wright, and Margaret Hudson. 2005. "Partnering 'out' and Fitting in: Residential Segregation and the Neighbourhood Contexts of Mixed-Race Households." *Population, Space and Place* 11: 299–324.
- Iceland, John and Gregory Sharp. 2013. "White Residential Segregation in U.S. Metropolitan Areas: Conceptual Issues, Patterns, and Trends from the U.S. Census, 1980 to 2010." *Population Research and Policy Review* pp. 1–24.
- Iceland, John, Gregory Sharp, and Jeffrey M. Timberlake. 2013. "Sun Belt Rising: Regional Population Change and the Decline in Black Residential Segregation, 1970–2009." *Demography* 50: 97–123.
- Iceland, John, Cicely Sharpe, and Erika Steinmetz. 2005. "Class differences in African American residential patterns in US metropolitan areas: 1990–2000." *Social Science Research* 34: 252–266.
- Iceland, John and Rima Wilkes. 2006. "Does Socioeconomic Status Matter? Race, Class, and Residential Segregation." *Social Problems* 53: 248–273.
- Ihlanfeldt, Keith R. and Benjamin Scafidi. 2002. "Black Self-Segregation as a Cause of Housing Segregation: Evidence from the Multi-City Study of Urban Inequality." *Journal of Urban Economics* 51: 366–390.
- Ihlanfeldt, Keith R. and Benjamin Scafidi. 2004. "Whites' neighbourhood racial preferences and neighbourhood racial composition in the United States: evidence from the multi-city study of urban inequality." *Housing Studies* 19: 325.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75: 83–119.
- Krysan, Maria. 2002. "Whites Who Say They'd Flee: Who Are They, and Why Would They Leave?" *Demography* 39: 675–696.
- Krysan, Maria and Michael D. M. Bader. 2009. "Racial Blind Spots: Black-White-Latino Differences in Community Knowledge." *Social Problems* 56: 677–701.
- Krysan, Maria, Mick P. Couper, Reynolds Farley, and Tyrone A. Forman. 2009. "Does Race Matter in Neighborhood Preferences? Results from a Video Experiment." *American Journal of Sociology* 115: 527–559.
- Krysan, Maria and Reynolds Farley. 2002. "The Residential Preferences of Blacks: Do They Explain Persistent Segregation?" *Social Forces* 80: 937–980.

- Laurie, Alexander J. and Narendra K. Jaggi. 2003. "Role of 'Vision' in Neighbourhood Racial Segregation: A Variant of the Schelling Segregation Model." *Urban Studies* 40: 2687–2704.
- Lee, Brian H. Y. and Paul Waddell. 2010. "Residential Mobility and Location Choice: A Nested Logit Model with Sampling of Alternatives." *Transportation* 37: 587–601.
- Lewis, Valerie A., Michael O. Emerson, and Stephen L. Klineberg. 2011. "Who We'll Live With: Neighborhood Racial Composition Preferences of Whites, Blacks and Latinos." *Social Forces* 89: 1385–1407.
- Logan, John R. 1978. "Growth, Politics, and the Stratification of Places." *American Journal of Sociology* 84: 404–416.
- Logan, John R. 2013. "The Persistence of Segregation in the 21st Century Metropolis." *City & Community* 12: 160–168.
- Logan, John R., Richard D. Alba, and Shu-Yin Leung. 1996. "Minority Access to White Suburbs: A Multiregional Comparison." *Social Forces* 74: 851–881.
- Logan, John R., Brian J. Stults, and Reynolds Farley. 2004. "Segregation of Minorities in the Metropolis: Two Decades of Change." *Demography* 41: 1–22.
- Louviere, Jordan J., David A. Hensher, and Joffre Dan Swait. 2000. *Stated Choice Methods: Analysis and Applications*. Cambridge University Press.
- MacDonald, John S. and Leatrice D. MacDonald. 1964. "Chain Migration Ethnic Neighborhood Formation and Social Networks." *The Milbank Memorial Fund Quarterly* 42: 82.
- Macy, Michael W. and Arnout Van De Rijt. 2006. "Ethnic Preferences and Residential Segregation: Theoretical Explorations Beyond Detroit." *The Journal of Mathematical Sociology* 30: 275–288.
- Massey, Douglas S. 1985. "Ethnic Residential Segregation: A Theoretical Synthesis and Empirical Review." *Sociology and Social Research* 69: 315–350.
- Massey, Douglas S. 2005. "Racial Discrimination in Housing: A Moving Target." *Social Problems* 52: 148–151.
- Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19: 431–466.
- Massey, Douglas S. and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67: 281–315.
- Massey, Douglas S. and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

- Massey, Douglas S. and Kristin E. Espinosa. 1997. "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102: 939–999.
- Massey, Douglas S. and Mary J. Fischer. 1999. "Does Rising Income Bring Integration? New Results for Blacks, Hispanics, and Asians in 1990." *Social Science Research* 28: 316–326.
- Massey, Douglas S. and Mary J. Fischer. 2000. "How Segregation Concentrates Poverty." *Ethnic and Racial Studies* 23: 670–691.
- McFadden, Daniel. 1978. "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, edited by Anders Karlqvist, Lars Lundqvist, Folke Snickers, and Jörgen W. Weibull, volume 3 of *Studies in Regional Science and Urban Economics*, pp. 75–96. North-Holland Publishing Company.
- McFadden, Daniel. 1980. "Econometric Models for Probabilistic Choice Among Products." *The Journal of Business* 53: S13–S29.
- McFadden, Daniel. 2001. "Economic Choices." *The American Economic Review* 91: 351–378.
- McFadden, Daniel and Kenneth Train. 2000. "Mixed MNL models for discrete response." *Journal of Applied Econometrics* 15: 447–470.
- Morikawa, Takayuki. 1989. *Incorporating Stated Preference Data in Travel Demand Analysis*. Doctoral thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Morikawa, Takayuki. 1994. "Correcting state dependence and serial correlation in the RP/SP combined estimation method." *Transportation* 21: 153–165.
- Ondrich, Jan, Alex Stricker, and John Yinger. 1999. "Do Landlords Discriminate? The Incidence and Causes of Racial Discrimination in Rental Housing Markets." *Journal of Housing Economics* 8: 185–204.
- Pais, Jeremy, Scott J. South, and Kyle Crowder. 2012. "Metropolitan Heterogeneity and Minority Neighborhood Attainment: Spatial Assimilation or Place Stratification?" *Social Problems* 59: 258–281.
- Peterson, Christene E., Anne R. Pebley, Narayan Sastry, Karen Yuhas, Bonnie Ghosh-Dastidar, Ann C. Haas, Jesse Gregory, and Marianne P. Bitler. 2012. "The Los Angeles Family and Neighborhood Survey, Wave 2 User's Guide and Codebook." Working Paper WR-240/20-LAFANS, RAND Corp., Santa Monica, CA.
- Peterson, Christine E., Narayan Sastry, Anne R. Pebley, Bonnie Ghosh-Dastidar, Stephanie Williamson, and Sandraluz Lara-Cinisomo. 2004. "The Los Angeles Family and Neighborhood Survey: Codebook."

- Pinjari, Abdul Rawoof, Ram M. Pendyala, Chandra R. Bhat, and Paul A. Waddell. 2011. "Modeling the Choice Continuum: An Integrated Model of Residential Location, Auto Ownership, Bicycle Ownership, and Commute Tour Mode Choice Decisions." *Transportation* 38: 933–958.
- Quillian, Lincoln. 1996. "Group Threat and Regional Change in Attitudes Toward African-Americans." *The American Journal of Sociology* 102: 816–860.
- Quillian, Lincoln. 1999. "Migration Patterns and the Growth of High-Poverty Neighborhoods, 1970-1990." *The American Journal of Sociology* 105: 1–37.
- Quillian, Lincoln. 2002. "Why Is Black-White Residential Segregation So Persistent?: Evidence on Three Theories from Migration Data." *Social Science Research* 31: 197–229.
- Quillian, Lincoln. 2012. "Segregation and Poverty Concentration The Role of Three Segregations." *American Sociological Review* 77: 354–379.
- Quillian, Lincoln. 2014. "Does Segregation Create Winners and Losers? Residential Segregation and Inequality in Educational Attainment." *Social Problems* 61: 402–426.
- Quillian, Lincoln and Devah Pager. 2001. "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime." *American Journal of Sociology* 107: 717–767.
- Reardon, Sean F. and Glenn Firebaugh. 2002. "Measures of Multigroup Segregation." *Sociological Methodology* 32: 33–67.
- Reardon, Sean F., Stephen A. Matthews, David O'Sullivan, Barrett A. Lee, Glenn Firebaugh, Chad R. Farrell, and Kendra Bischoff. 2008. "The Geographic Scale of Metropolitan Racial Segregation." *Demography* 45: 489–514.
- Reibel, Michael. 2000. "Geographic Variation in Mortgage Discrimination: Evidence from Los Angeles." *Urban Geography* 21: 45–60.
- Revelt, David and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics* 80: 647–657.
- Revelt, David and Kenneth Train. 2001. "Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier." *Econometrics* 0012001, EconWPA.
- Ross, Stephen L. and Margery Austin Turner. 2005. "Housing Discrimination in Metropolitan America: Explaining Changes between 1989 and 2000." *Social Problems* 52: 152–180.
- Sampson, Robert J. 2011. *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago ; London: The University of Chicago Press.
- Sampson, Robert J. and Stephen W. Raudenbush. 2004. "Seeing Disorder: Neighborhood Stigma and the Social Construction of 'Broken Windows'." *Social Psychology Quarterly* 67: 319–342.

- Sastry, Narayan, Bonnie Ghosh-Dastidar, John Adams, and Anne R. Pebley. 2006. "The Design of a Multilevel Survey of Children, Families, and Communities: The Los Angeles Family and Neighborhood Survey." *Social Science Research* 35: 1000–1024.
- Schelling, Thomas C. 1969. "Models of Segregation." *The American Economic Review* 59: 488–493.
- Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *The Journal of Mathematical Sociology* 1: 143–186.
- Sharkey, Patrick. 2012. "Temporary Integration, Resilient Inequality: Race and Neighborhood Change in the Transition to Adulthood." *Demography* 49: 889–912.
- Singh, Abhinav, Dmitri Vainchtein, and Howard Weiss. 2009. "Schelling's Segregation Model: Parameters, Scaling, and Aggregation." *Demographic Research* 21: 341–366.
- Sjaastad, Larry A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy* 70: 80–93.
- South, Scott J., Kyle Crowder, and Erick Chavez. 2005a. "Exiting and Entering High-Poverty Neighborhoods: Latinos, Blacks and Anglos Compared." *Social Forces* 84: 873–900.
- South, Scott J., Kyle Crowder, and Erick Chavez. 2005b. "Migration and Spatial Assimilation among U.S. Latinos: Classical versus Segmented Trajectories." *Demography* 42: 497–521.
- South, Scott J., Kyle Crowder, and Jeremy Pais. 2008. "Inter-Neighborhood Migration and Spatial Assimilation in a Multi-Ethnic World: Comparing Latinos, Blacks and Anglos." *Social Forces* 87: 415–443.
- South, Scott J., Kyle Crowder, and Jeremy Pais. 2011. "Metropolitan Structure and Neighborhood Attainment: Exploring Intermetropolitan Variation in Racial Residential Segregation." *Demography* 48: 1263–1292.
- South, Scott J. and Kyle D. Crowder. 1997. "Escaping Distressed Neighborhoods: Individual, Community, and Metropolitan Influences." *American Journal of Sociology* 102: 1040–1084.
- South, Scott J. and Kyle D. Crowder. 1998a. "Housing Discrimination and Residential Mobility: Impacts for Blacks and Whites." *Population Research and Policy Review* 17: 369–387.
- South, Scott J. and Kyle D. Crowder. 1998b. "Leaving the 'Hood: Residential Mobility between Black, White, and Integrated Neighborhoods." *American Sociological Review* 63: 17–26.
- Stark, Oded and David E. Bloom. 1985. "The New Economics of Labor Migration." *The American Economic Review* 75: 173–178.

- StataCorp. 2011. *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.
- Statistics Sweden. 2014. *Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA)*. Stockholm: Statistics Sweden.
- Swait, Joffre, Jordan J. Louviere, and Michael Williams. 1994. "A sequential approach to exploiting the combined strengths of SP and RP data: Application to freight shipper choice." *Transportation* 21: 135–152.
- Taeuber, Karl E. and Alma F. Taeuber. 1965. *Negroes in Cities: Residential Segregation and Neighborhood Change*. Chicago: Aldine Publishing Company.
- Taeuber, Karl E. and Alma F. Taeuber. 1976. "A Practitioner's Perspective on the Index of Dissimilarity." *American Sociological Review* 41: 884–889.
- Theil, Henri. 1972. *Statistical Decomposition Analysis with Applications in the Social and Administrative Sciences*. Amsterdam: North-Holland.
- Theil, Henri and Anthony J. Finizza. 1971. "A Note on the Measurement of Racial Integration of Schools by Means of Informational Concepts." *The Journal of Mathematical Sociology* 1: 187–193.
- Thernstrom, Stephan and Abigail Thernstrom. 1997. *America in Black and White: One Nation, Indivisible*. New York, NY: Simon & Schuster.
- Timberlake, Jeffrey M. and John Iceland. 2007. "Change in Racial and Ethnic Residential Inequality in American Cities, 1970–2000." *City & Community* 6: 335–365.
- Train, Kenneth. 1986. *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*. Number 10 in MIT Press series in transportation studies. Cambridge, Mass: MIT Press.
- Train, Kenneth. 2000. "Halton Sequences for Mixed Logit." *UC Berkeley Department of Economic Working Paper qt6zs694tp*.
- Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. Cambridge ; New York: Cambridge University Press, 2nd ed edition.
- Train, Kenneth E. 2008. "EM algorithms for nonparametric estimation of mixing distributions." *Journal of Choice Modelling* 1: 40–69.
- Treiman, Donald J and Hye-Kyung Lee. 1996. "Income Differences among 31 Ethnic Groups in Los Angeles." In *Social Differentiation and Social Inequality: Essays in Honor of John Pock*, edited by James N. Baron, David B. Grusky, and Donald J. Treiman, Social inequality series, pp. 37–82. Boulder, CO: Westview Press.
- Turner, Margery Austin, Stephen L. Ross, George C. Galster, and John Yinger. 2002. "Discrimination in Metropolitan Housing Markets: National Results from Phase I HDS 2000." Washington, D.C.: The Urban Institute.

- Turner, Margery Austin, Rob Santos, Diane K. Levy, Doug Wissoker, Claudia Aranda, Rob Pitingolo, and Urban Institute. 2013. "Housing Discrimination against Racial and Ethnic Minorities 2012: Full Report." Washington D.C.: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.
- Van De Rijt, Arnout, David Siegel, and Michael Macy. 2009. "Neighborhood Chance and Neighborhood Change: A Comment on Bruch and Mare." *American Journal of Sociology* 114: 1166–1180.
- Waddell, Paul, Chandra Bhat, Naveen Eluru, Liming Wang, and Ram Pendyala. 2007. "Modeling Interdependence in Household Residence and Workplace Choices." *Transportation Research Record: Journal of the Transportation Research Board* 2003: 84–92.
- Wagmiller, Robert L., Mary Clare Lennon, Li Kuang, Philip M. Alberti, and J. Lawrence Aber. 2006. "The Dynamics of Economic Disadvantage and Children's Life Chances." *American Sociological Review* 71: 847–866.
- Williams, Richard, Reynold Nesiba, and Eileen Diaz McConnell. 2005. "The Changing Face of Inequality in Home Mortgage Lending." *Social Problems* 52: 181–208.
- Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Winship, Christopher. 1977. "A Revaluation of Indexes of Residential Segregation." *Social Forces* 55: 1058–1066.
- Wodtke, Geoffrey T., David J. Harding, and Felix Elwert. 2011. "Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation." *American Sociological Review* 76: 713–736.
- Xie, Yu and Xiang Zhou. 2012. "Modeling individual-level heterogeneity in racial residential segregation." *Proceedings of the National Academy of Sciences* 109: 11646–11651.
- Yinger, John. 1995. *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination*. New York: Russell Sage Foundation.
- Zhang, Junfu. 2004. "A Dynamic Model of Residential Segregation." *The Journal of Mathematical Sociology* 28: 147–170.
- Zhou, Min, Jennifer Lee, Jody Agius Vallejo, Rosaura Tafoya-Estrada, and Yang Sao Xiong. 2008. "Success Attained, Deterred, and Denied: Divergent Pathways to Social Mobility in Los Angeles's New Second Generation." *The ANNALS of the American Academy of Political and Social Science* 620: 37–61.
- Zubrinsky, Camille L. and Lawrence Bobo. 1996. "Prismatic Metropolis: Race and Residential Segregation in the City of the Angels." *Social Science Research* 25: 335–374.