SIMULATING THE SOCIAL DYNAMICS OF SPATIAL DISPARITY THROUGH NEIGHBORHOOD NETWORK EVOLUTION

BY

SARA S. METCALF

B.S., Texas A&M University, 1996
M.S., Massachusetts Institute of Technology, 2001

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Geography
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2007

Urbana, Illinois

ABSTRACT

The concept of spatial disparity is so familiar that we often don't recognize the problems implicit in referencing the "wrong side of town" in casual conversation. But for communities expanding in area while shrinking or stagnating in population, this problem becomes even more obvious in the dichotomy between "new" and "old" or "rich" and "poor" parts of town. And the "wrong side of town" becomes even easier to avoid. While this problem of disparity is not new, neighborhoods in flux still struggle to avoid a fate of either deterioration or displacement.

In research on spatial disparity, the relationship between social ties and location remains open territory. While local interactions have been studied, they have not been linked to aggregate patterns of community cohesiveness or fragmentation. Recent developments in computer simulation techniques now enable explicit representation of local community ties as they evolve over time. These developments allow systematic exploration of relationships between individual choices about where to live and social network structure, and between neighborhood and community networks.

Following an overview of enabling theories and developments, I develop a spatial dynamic simulation model that enables virtual experimentation with both friendship and location choice. In the model, individuals interact over space and time in an agent-based framework, and neighborhood networks emerge from these simulated interactions. I calibrate the model to the community of Danville, Illinois, which has experienced an increase in fringe development in the absence of population growth. Heuristics for modeling individual choices draw upon field interviews with residents of a rapidly changing Danville neighborhood. Model parameters for moving behavior, individual attributes, and network structure are estimated from a variety of empirical sources. Alternative model structures are calibrated to observed migration patterns to discern the relative role of social influence on neighborhood choice. With the calibrated model, policy scenarios may be explored to help communities avoid a fate of fragmentation.

The significance of this project is at once substantive, in addressing the persistence of spatial disparity in fragmenting communities, and also methodological, in bridging research domains that have not yet been functionally linked.

ACKNOWLEDGMENTS

I am most indebted to my committee for their patience and support during this time. My advisor, Bruce Hannon, always provided intellectual support, challenging me but leaving plenty of room to explore new territory on my own. Mark Paich helped me navigate my way back to graduate school, and has always been able to see future possibilities clearly. Sara McLafferty has been a reliable source of direct and constructive feedback. Mu, Lan provided quick responses and efficient solutions to my sporadic GIS queries. Jim Westervelt shared modeling expertise and also helped me see the perspective of my committee.

I am grateful for several technical consultations with Lyle Wallis on the use of Java for pre-processing data. Conversations with Alfred Hubler and Mike Woodley helped me to structure my approach. As a study partner, Veronica Escamilla enabled me to maintain momentum during the crux of my data processing tasks. Geography colleagues David Wilson, Shane Csiki, Liz Lyon and Rob Whiting provided early feedback on the proposal for this research. Thanks to Lauren Fontenot Beck and Chris Wilcock for help in depositing the thesis.

Project support from NSF Award HSD-0433165 has been instrumental to this work. Thanks also to the LEAM (Land-use Evolution and impact Assessment Model) project for exposing me to the role of spatial disparity in perpetuating urban sprawl. Conversations with Varkki George Pallathucheril in particular helped to frame the social dynamics of this problem.

The residents of Danville were very welcoming of this research project. Members of the Kentucky-Tennessee-Delaware neighborhood association were particularly open with me, both during formal and informal interviews and observations. City officials Chris Milliken, John Dreher, and John Heckler helped me to get started and to augment the interviews with quantitative data. Brad and Sarah Boesdorfer helped to extend my own network into Danville.

Several networks of women have helped sustain my passion for this project: the cooking group, the Thursday afternoon group, and the Women Leading Sustainability. Many thanks also go to my extended family for their support during the long process of “finishing” the dissertation. My father, Allan Metcalf, helped me articulate my abstract using only one-syllable words. My Texas family—Alex, Teri, Bob, Mike, Fang, Itchy, Tiny and Sunshine—provided daily encouragement during the final phase of this project. My mother, Teri Metcalf, has helped me most of all by simply showing me how to be.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION ... 1
- 1.1 Motivation ... 1
- 1.2 Problem Statement ... 2
- 1.3 Research Hypothesis ... 4
- 1.4 Project Significance ... 6
- 1.5 Research Design ... 6
  - 1.5.1 Study Site ... 8
  - 1.5.2 Model Development ... 9
  - 1.5.3 Model Calibration ... 9
- 1.6 Virtual Experimentation ... 10

CHAPTER 2. THEORY ... 12
- 2.1 Agent-Based Modeling ... 12
  - 2.1.1 An Emerging Methodology ... 14
  - 2.1.2 Feedback and Iteration ... 16
- 2.2 Social Networks ... 18
  - 2.2.1 Social Networks in Geography ... 19
  - 2.2.2 Dynamic Social Networks ... 21
- 2.3 Conceptual Integration ... 23
  - 2.3.1 Learning is Social ... 23
  - 2.3.2 Social Networks are Spatial ... 25
  - 2.3.3 Neighborhood Choice ... 26

CHAPTER 3. EMPIRICAL CONTEXT ... 28
- 3.1 Study Site ... 28
  - 3.1.1 Recent Trends ... 29
  - 3.1.2 Recent Studies ... 31
- 3.2 Data Sources ... 32
  - 3.2.1 Qualitative Data ... 33
  - 3.2.2 Quantitative Data ... 34
- 3.3 Data Processing ... 34
  - 3.3.1 Interview Heuristics ... 34
  - 3.3.2 Parcel Data Processing ... 37
  - 3.3.3 Merging Owner Records ... 39
  - 3.3.4 Census Data Extraction ... 40
- 3.4 Analysis of Geographic Patterns ... 41
  - 3.4.1 Distance Between Households Sharing Last Name ... 41
  - 3.4.2 Creating the Owner Migration Pattern ... 46

CHAPTER 4. MODEL DEVELOPMENT ... 50
- 4.1 Model Conceptualization ... 50
- 4.2 CaseTown Model Structure ... 53
  - 4.2.1 Creating Model Objects from Data ... 54
  - 4.2.2 AnyLogic Model Structure ... 56

4.3 Model Dynamics ... 58
4.3.1 Stock and Flow Representation of System Dynamics ... 59
4.3.2 Decision Rules ... 60
4.4 Simulating Spatial Disparity ... 64

CHAPTER 5. CALIBRATION RESULTS ... 67
5.1 Specifying Alternatives ... 67
5.2 Strategy ... 69
5.3 Testing ... 71
5.3.1 Solution Space ... 72
5.3.2 Regression Results ... 78

CHAPTER 6. CONCLUSIONS ... 82
6.1 Insights ... 83
6.2 Model Extensions ... 84
6.3 Policy Analyses ... 85
6.4 Implications ... 86

REFERENCES ... 88

APPENDICES ... 93
Appendix A: Interview Guide ... 94
Appendix B: Neighborhood Research Report-Out ... 97
Appendix C: Metadata for Parcel Dataset ... 104
Appendix D: Example Raw (Unstructured) Data Files for Model Objects ... 105
Appendix E: Observed Migration Patterns ... 109
Appendix F: Java Classes for Data Processing ... 111
Processing Package ... 111
Structure Package ... 116
CaseTown Package ... 118
Appendix G: AnyLogic Model Structure ... 121
Main Active Object Class ... 121
Initialize Active Object Class ... 125
Household Active Object Class ... 126
Appendix H: System Dynamics Basics ... 131

List of Figures

Figure 1. Relationships between Social Networks and Neighborhood Choice 5
Figure 2. Learning from the Virtual Experiments of Modeling. 11
Figure 3. Agent-Based Modeling Framework. 13
Figure 4. Conceptualizing Agents as Part of Stocks and Flows 15
Figure 5. The Iterative Process of Pattern-Oriented Modeling. 16
Figure 6. Feedback Across Scale in Agent-Based Models 17
Figure 7. Example Small-World and Scale-Free Networks.. 22
Figure 8. Visualization of Interactive Learning.. 24
Figure 9. Relative Location of Study Site: Danville, Illinois, USA. 28
Figure 10. Trends in Population and Land Area for Danville, IL, USA...................................... 29
Figure 11. Danville Median Income by BlockGroup 30
Figure 12. Map of Danville Wards 32
Figure 13. Selected Danville Block Boundaries 37
Figure 14. Difficulty Associating Parcels with Census Blocks 38
Figure 15. Distribution of Distance Between Households... 43
Figure 16. Distance Between Households by Number of Households Sharing Last Names 44
Figure 17. Distribution of Unique Names by Number of Households Sharing Last Names........ 45
Figure 18. Migration Within Danville from 2001 to 2003 46
Figure 19. Ownership Changes From 2001 to 2003 Within Danville 47
Figure 20. Overall Ownership Changes From 2001 to 2003 48
Figure 21. Example Template for a Spatial Social Network 51
Figure 22. Screenshot in the AnyLogic Simulation Environment 52
Figure 23. Map of Owner Occupancy in Danville Census Blocks 54
Figure 24. Process for Creating Model Objects 55
Figure 25. Model Structure Screenshot in AnyLogic 57
Figure 26. Causal Map of Neighborhood Dynamics 59
Figure 27. Transitions Between Rent and Own Status 63
Figure 28. Gini Coefficient and Fraction of Local Ties over Years of a Single Simulation 65
Figure 29. Model Structure Alternatives 67
Figure 30. Calibration Strategy.. 70
Figure 31. View of Optimization Process.. 71
Figure 32. Total Move Error and Directional (BlockGroup) Error for 1000 Experiments 73
Figure 33. Variation of Move Error by Amount of Error Among Filtered Solutions 74
Figure 34. Filtered Patterns of Parameter Effect Ratios 75
Figure 35. Filtered Comparison of Spatial and Social Influences on Choice 76
Figure 36. Cluster Analysis of Parameters in "Good" Solution Space 77
Figure 37. Screenshot of Main Active Object Class Structure 121
Figure 38. Screenshot of Household Active Object Class Structure 126
Figure 39. Example of Reinforcing Feedback Loop.. 131
Figure 40. Example of Balancing Feedback Loop... 132
Figure 41. Example of Interaction of Reinforcing and Balancing Effects 132
Figure 42. Stock and Flow Representation of Bathtub Dynamics... 133

List of Tables

Table 1. Outline of Research Protocol 7
Table 2. Qualitative and Quantitative Data Collected for Danville 33
Table 3. Example of Owner Matching Across Years 40
Table 4. Comparison of Households with Shared and Unique Last Names 42
Table 5. Description of Parameters Varied to Create Alternative Model Structures 68
Table 6. Regression Summary of Parameter Effects on Average Move Error 78
Table 7. Regression of Model Parameters (Linear Form) on Average Move Error 79
Table 8. Regression of Model Parameters (with Squared Terms) on Average Move Error 80
Table 9. Row Excerpt from Parcel Data (13166parcels.csv) 105
Table 10. Row Excerpt from Owner Data File (7576owners.csv) 106
Table 11. Row Excerpt from Block Data File (764blocks.csv) 107
Table 12. Column Excerpt from BlockGroup Data File (28blockGroups.csv) 108
Table 13. Owner Migration by BlockGroup from 2001 to 2003 109
Table 14. Owner Migration by BlockGroup from 2003 to 2005 110

List of Equations

Equation 1. Social Network Choice 61
Equation 2. Neighborhood Choice 62
Equation 3. Formula for Measuring Gini Coefficient 64

CHAPTER 1. INTRODUCTION

This thesis applies novel simulation techniques to the persistent problem of spatial socioeconomic disparity. This chapter introduces the problem and the research design. The following chapters then outline the enabling theoretical developments for this work, the empirical context, and the model development and calibration results. By using observed migration patterns to compare with simulation results, this research tests alternative model structures of the recursive influence of social networks on neighborhood choice. In short, this research steps through one full iteration of the modeling cycle and provides a foundation for future work on operationalizing social theories of spatial disparity.

1.1 Motivation

The initial conception of this project began to take shape when I started research at the University of Illinois in the late spring of 2003. I was hired as a research assistant for the Land-use Evolution and impact Assessment Model (LEAM), a multidisciplinary and multi-year grant to simulate the drivers and impacts of urban sprawl. At its core, LEAM draws on economic drivers to project population growth for the region and distributes the population spatially by travel time proximity to attractors, such as city amenities, major roads, and recreational areas. For the St. Louis region, if undeveloped land exists close to the city, it would have a high probability of being developed in the model. The principals at LEAM were concerned that this did not capture the realities of abandoned structures and spaces in some of these areas. Particularly notorious was East St. Louis, just across the Mississippi river in the floodplain. Its history has been marked by industrial investment and subsequent abandonment, and its residents today are nearly all African-American.

At a May 2003 workshop in the case study region of St. Louis, over 100 stakeholders (e.g., planners, activists, ministers, mayors) brainstormed drivers and impacts of urban sprawl. Social factors were frequently mentioned as drivers – such as school quality, fear of crime, safety, lifestyle, class and race segregation – with 27% of stakeholder votes of significance, but social factors were rarely mentioned as scenarios, garnering only 5% of stakeholder votes for consideration. Recognition of social phenomena was key to understanding regional development, but appeared to be outside the domain of leverage for most stakeholder participants. This gap in representation of social dilemmas, alongside the model's inability to simulate sprawl alongside population decline, suggested that the dynamics of urban decline would benefit from testing the underlying social dynamics. I wanted to understand how it was that some places became "undesirable" over time. My intuition was that social influences shape individual perceptions and choices of where to live, and could therefore reinforce outmigration tendencies such as "white flight."[1]

My hypothesis of social influence on neighborhood development begins with my own experience. Growing up less than two hours from St. Louis, the spatial wisdom conveyed to me was "Never get lost in East St. Louis." I do not recall which parent, teacher or peer first conveyed this to me, but family and friends reinforced the message on trips to St. Louis (especially once I reached driving age). And so I avoided East St. Louis – until 2003 as part of the East St. Louis Action Research Project (ESLARP). Instead of the fear and danger I anticipated, what I experienced was a sense of loss, struggle, and abandonment. In many ways, East St. Louis felt similar to many places in my own hometown of Jacksonville, Illinois. Jacksonville had its own "wrong side of town" that was to be avoided, and it's remarkable to me in hindsight how well segregated along socioeconomic and starkly racial lines a relatively small town could be. Even more amazing to me is recognizing my own role in perpetuating such segregation via my own avoidance of undesirable places. My shock is not around the legitimacy of fear, but around the lack of concern that perpetual avoidance can produce. I grew up essentially ignorant of two undesirable places: East St. Louis and the wrong side of the tracks in my own hometown. But it's not just these particular places; it's the individual acceptance of places as off-limits that I find problematic for coming toward any semblance of tolerance or equity in our society.

1.2 Problem Statement

Even in the absence of population growth, many communities continue to experience urban sprawl, or low-density fringe development (Downs 2000). Since a household's ability to move to a new home usually requires an income well above the poverty level, new development serves to separate those who can afford to choose from those who cannot. In this way, sprawl

[1] "White flight" is the phrase ascribed to the neighborhood succession that manifested racially beginning in the 1950s, in which whites moved out of middle-class neighborhoods as blacks moved in.

dynamics exacerbate spatial disparity between socioeconomic classes. But underlying the aggregate phenomenon of sprawl is the individual choice of where to live. For a current resident, this is a choice of whether to stay or leave a neighborhood in favor of another neighborhood or another community altogether. For a newcomer, this is a choice of where to settle upon arrival. In aggregate these choices shape the spatial and temporal dimensions of urban dynamics such as sprawl and its corollary, spatial disparity.

The central problem of this research project is the social dilemma of spatial disparity exhibited in urban areas. Spatial disparity, measured by broad income differences between distinct neighborhoods, is evidenced and reinforced by the phenomenon of urban sprawl (Downs 1999). This research project develops dynamics of spatial disparity from household decisions of where to live.

Choosing to move to a new neighborhood necessarily involves the decision to leave one's current neighborhood. The question of where we move to may be studied in terms of attractiveness due to proximity to amenities, aesthetics, and the like. However, the role of social ties in determining neighborhood desirability remains largely unaddressed. Moreover, these ties influence perceptions of places as undesirable. Addressing the role of these apparent intangibles may shed light on the mechanisms of phenomena such as "white flight," which remain opaque relative to standard models of neighborhood utility. While "white flight" may be an extreme dynamic ignited by racism, it highlights the importance of social influence on perceptions that in turn relate to neighborhood choice.

To systematically assess the role of social influence in neighborhood choice, this research simulates these facets of neighborhood choice with computerized representations of dynamic social networks. A social network is an abstraction of the interactions between people that define relationships (a.k.a. "social ties"). Functionally, these relationships create trusted channels for communication. In the community context, social networks encompass family, friends, and advisors.

The hypothesized recursive effect of social networks on neighborhood choice becomes difficult to test due to the inherent impracticality of experimenting with policies on a real community. Virtual experimentation via computer models is therefore a logical if not essential way to explore the effects of social networks on neighborhood choice over time. While a computer model will not predict the future of a real community, it is the only practical means of testing alternative assumptions and policies in an internally consistent framework. And it is this process of modeling and testing, this virtual experimentation, which offers opportunities for insight or surprise relative to expectations.

1.3 Research Hypothesis

The dynamics of spatial disparity between socioeconomic groups are not adequately understood. Available theories have not explained this phenomenon sufficiently to help guide community policy to stabilize neighborhood dynamics. Therefore, more operational theories are needed for systematic analysis of spatial disparity. By operationalizing hypothesized relationships between driving forces in a simulation model, this research breaks ground for a theoretical framework to be built on transparent and internally consistent assumptions of spatiotemporal household choice.

Centered on the question of neighborhood choice, this study hypothesizes that social networks influence migration patterns that reinforce spatial disparity in urban communities. This general hypothesis may be stated in the following specific forms:

- H1. Social ties influence households' choice of whether to stay in, leave, or move to a neighborhood.
- H2. Social networks evolve over time as opportunities for interaction change (through proximity, income parity, and other factors).
- H3. The influence of social networks on neighborhood choice contributes to spatial disparity of socioeconomic status.

This research focuses on the explicit testing of H1, while providing a simulation framework to enable testing of H2 and H3. To test the H1 hypothesis of social influence on neighborhood choice, I develop a computer simulation tool for integrating the recursive influence of social networks on neighborhood choice over time. The community of Danville, Illinois is used as a study site to provide real-world relevance, through both qualitative interviews and quantitative empirical data. Such an integrated and applied simulation offers insight into the long-studied but still open dilemma of persistent inequality as evidenced through spatial disparity.

The purpose of this research is to create a more complete understanding of the role that social networks can play in shaping community structure through neighborhood choice. Figure 1 below illustrates the recursive relationship between social networks and neighborhood choice. Social networks influence individuals' choice of neighborhoods through communication of what is desirable, as well as opportunities to visit different neighborhoods. The presence of friends or family in one neighborhood may make such a neighborhood more appealing, whereas the absence of such social ties may discourage a move to a new location. In turn, neighborhoods create spaces for interaction and thus adjustment of social ties due to proximity. This simple feedback diagram illustrates how social networks shape neighborhood choice and are in turn shaped by such choice, insofar as the social network depends in part on neighborhood proximity. Additional influences may be represented as exogenous to the relationships shown.

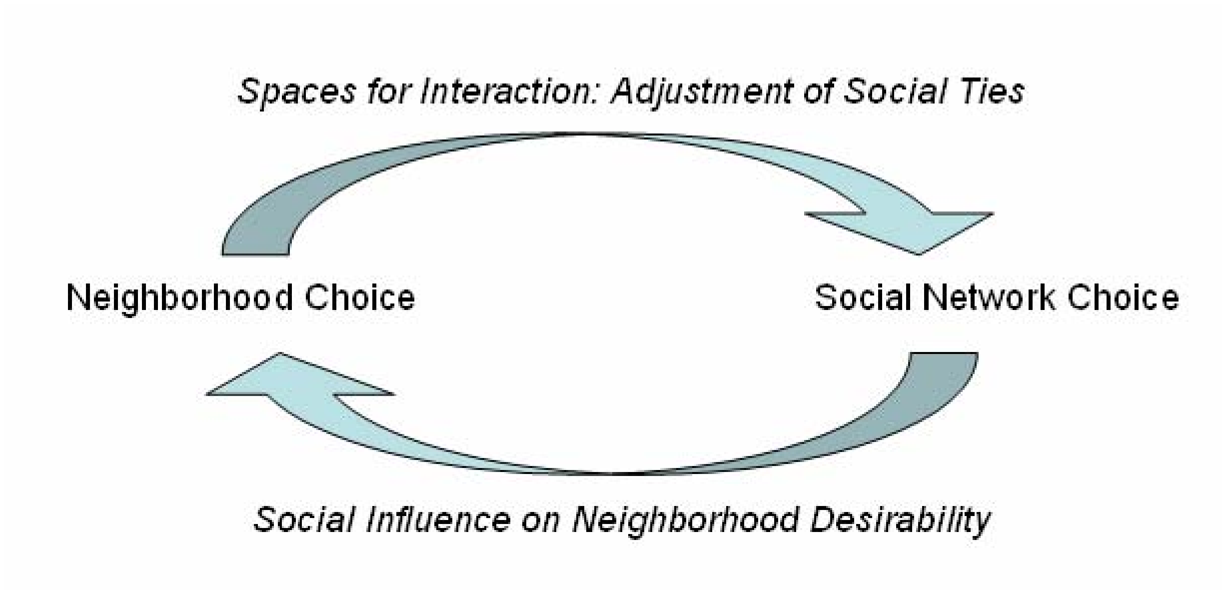

Figure 1. Relationships between Social Networks and Neighborhood Choice

It is worth noting that this research does not attempt to capture *complete* social networks for a community. Rather, the influence of these networks will be indirectly estimated from observed migration patterns. And while the symptom of sprawl has motivated the study, this research does not simulate sprawl per se. The research intent is simply to determine whether the sheer presence of social networks can inform understanding of neighborhood change and community fragmentation.

1.4 Project Significance

Increased community fragmentation is a concern for community efficacy in establishing new economic opportunities, and for potential conflict between sub-communities. A further concern is that fragmented communities reinforce their exclusive biases through intra-group communication. It is through such communication that social influence on neighborhood choice may create potent reinforcing dynamics of disparity. This research addresses the question of community fragmentation through the use of simulated spatial social networks calibrated to the case of Danville, Illinois. This simulation framework enables systematic evaluation of policies that affect the trajectory of community fragmentation.

In addition to the central contribution of a coherent tool for policy evaluation, this study bridges a variety of specific research gaps:

- Explicit connections between social networks and neighborhood choice
- Evolution of social networks in spatially-explicit environments
- Calibration and evaluation of dynamic model structures using GIS data
- Integration of qualitatively elicited interviews with quantitative simulation techniques

The merit of the research consists of its integration of diverse approaches to social science, combining fieldwork and computer simulation to better understand how social networks evolve in the community context, and to visualize how such evolution may play out in terms of neighborhood choice over time. The broader impact of the project is that it addresses the persistence of spatial disparity that plagues policy makers and researchers. This research is expected to provide insight into areas of leverage for improving conditions of disparity in fragmenting urban areas.

1.5 Research Design

The design for this project involved data collection, model development, and calibration. Data collection methods for the city of Danville included participant observation and in-depth interviews with residents of a new neighborhood association in a low-to-middle income part of town that has been going through rapid changes in occupancy. The interviews were used to guide assumptions and provide ground-truth for the model, revealing the limitations of its scope. Geographic data were also obtained both from the city of Danville and the census at various scales. Owner address data for the years 2001, 2003, and 2005 were obtained from the city and used to derive migration patterns.

Alongside data collection was the model development, beginning with conceptualization and development of dynamic algorithms for household choice of social networks and neighborhoods. Model calibration brought these activities together to make the model relevant to the case of Danville. Calibration included identification of specific parameters based directly upon data, as well as indirect estimation of unobserved effects using broad migration patterns.

These methods link together disparate modeling paradigms. Geographic information systems (GIS), used to integrate the spatial data at various scales, are powerful but primarily static tools for spatial analysis. Social network analysis, used to measure properties of the simulated networks to assess cohesiveness, involves structural analysis of networks as different kinds of spaces, but does little to inform the dynamics. The third paradigm is agent-based modeling, the engine used to evolve the neighborhood networks. Agent-based modeling starts with the idea of an agent that may be mobile and updates its state (and perhaps the state of other entities) based upon decision rules or algorithms. Agents are thus discrete decision entities. The agent-based approach enables one-to-one mapping of the model to the system of interest, without requiring aggregation.

Table 1. Outline of Research Protocol

<table>
<tr><td>Data Collection:
1. Fieldwork
a. Participant Selection
b. In-depth Interviews
c. Content Analysis
2. GIS analysis
a. Parcel identification
b. Census blocks and blockgroups</td><td>Model Development:
1. Prototype template
a. Abstract neighborhoods
b. Spatial selection preferences
c. Socioeconomic selection
d. Update network as people move
2. Initial sensitivity testing
3. Alternative structure articulation</td></tr>
<tr><td colspan="2">Calibration:
1. Application of model template to Danville case
a. Parcel location; block-level occupancy and distance
b. Income distribution at blockgroup level
c. Choice heuristics from interviews
2. Determination of model structures that best fit observed patterns
3. Sensitivity testing and analysis of best fit structures</td></tr>
</table>

The research protocol is outlined in Table 1. As indicated in the protocol, data collection and initial model development activities were conducted in parallel. Such parallel development enabled transfer of insights between activities prior to integration. The data collection began with qualitative fieldwork for the case of Danville, Illinois. Such fieldwork served both to "ground truth" the simulation and also provide possible heuristics for social influence on neighborhood choice. In addition, GIS analysis was conducted for available parcel data juxtaposed with census area attributes at the block and blockgroup level. In parallel, the model development began with an abstract two-neighborhood template, with internalized probability of "rewiring" social connections based upon spatial and socioeconomic proximity. Sensitivity testing of the prototype model examined the effect of varying move probabilities (between neighborhoods and into or out of the community altogether) on overall community coherence, to what tipping points may exist in triggering fragmentation.

After testing the prototype and assimilating empirical data, alternative model structures were articulated to explore distinct effects on social networks and neighborhood choice. Where possible, such structures were achieved by constraining appropriate parameters in the utility formulations that govern choice. This model template was then applied to the Danville case, where individuals were allocated in space according to housing locations derived from GIS data. Household attributes were calibrated based upon census distributions at the block and blockgroup level. Moreover, choice heuristics were checked with interview insights for validity of assumptions. With the revised model template, alternative model structures were tested for the best fist to observed migration patterns in the city of Danville. Sensitivity testing and policy analysis of best fit structures provided the concluding insights of this project.

1.5.1 Study Site

Critical to the testing of the models applied to a real-world community was the inclusion of relevant data, both to directly calibrate and also to serve as a broader check on the underlying assumptions. This study utilized both qualitative interviews and quantitative GIS data derived from the census and the city of Danville. While far from isolated, the town of Danville, Illinois provided a coherent study site for fieldwork. Danville is a post-industrial community of about 34,000 people that has experienced population decline and stagnation, but still experiences sprawl in the form of new fringe development. As described in Chapter 3, Danville offers a glimpse of coherent community dynamics in both designating and dealing with undesirable spaces. The utility of the model for virtual experimentation may also be examined with scenarios relevant to decisions and concerns faced by the Danville community.

1.5.2 Model Development

During model development, alternative social network constructions were produced using different preferences for friendship formation. These networks then evolved with neighborhood and community migration, and may be tested for tipping points of fragmentation in the face of alternative migration probabilities. These social networks were spatially explicit, where one's location in two-dimensional Euclidean space connotes one's place of residence. The probability of "rewiring" one's social connections as inspired by small world dynamics (Watts 1999a) became endogenous to the system, depending on socioeconomic status similarity and neighborhood proximity as well as a fair degree of probability. Sensitivity testing was conducted with varying migration probabilities under this endogenous rewiring to explore the relative cohesiveness of the emergent community networks. The development of this abstract model was then extended to the real-world case community of Danville, Illinois.

The socioeconomic and spatial assumptions underlying network connectivity were derived from census data as integrated into the model with Geographic Information Systems (GIS). While GIS census attributes were available for areas such as census blocks and blockgroups, such data were disaggregated via probability distributions to correspond to individual households. Locations of individual households were determined by parcel data, which was available in GIS form by the city of Danville. A significant part of model development was the integration of data from diverse sources, as described in Chapters 3 and 4.

1.5.3 Model Calibration

As individual household objects were located in space via parcel data, they retained identification with census blocks and block groups. Such identification enabled heterogeneous parameterization of attributes based upon distribution of counts recorded in the census (e.g., 10 households with income less than $20,000, 5 houses unoccupied, etc.). Regardless of exact point location, accuracy of assignment was achieved only at the broader level of census area. But such accuracy was not of primary concern – rather the range of heterogeneity was considered to be of greater importance in structuring social networks. The translation into networks was guided by techniques of indirect estimation using socioeconomic data as developed by Conley and Topa (2003). In analyzing unemployment, Conley and Topa (2003) used broad spatial proximity to structure the social network of local interactions, with the majority occurring within and adjacent to the tract in question. While this strictly spatial algorithm was highly simplified, they demonstrated that it works better for anticipating shifts in unemployment than an aggregate black box approach would. In addition to the spatial considerations of social network structure, I incorporated social connections also based upon broad socioeconomic status. The specific attributes and algorithms for such connections were then varied for scenario analysis.

Development of the calibrated and applied model structures involved extensive testing – evaluating a range of parameters to investigate the effect on observed migration patterns. These calibration techniques are described in Chapter 5. In addition, the model was informally calibrated with insights gleaned from interviews with residents. Such interviews informed as well as revealed the limitations of the model relative to individual experience.

1.6 Virtual Experimentation

The goal of this research was not to produce one defining model but rather to provide an integrative simulation tool for considering alternative assumptions and theories about social networks as they relate to neighborhood choice. The importance of virtual experimentation with computer simulation models was central to the research design. Figure 2 highlights the role that such experimentation plays in influencing researchers' own mental models of the problem at hand. Alternative model structures are constructed and tested for insight. It is through the process of iteration in model design that generates insight about the question at hand. As emphasized in the problem statement, virtual experimentation enables studies of human behavior to be tested in ways that are not feasible or ethical in real communities. As a corollary, virtual experimentation offers a range of simulated realities that may or may not emerge in the "real world."

In Figure 2, this learning from virtual experimentation is illustrated as embedded as an iterative modeling process within learning from the real world. Observations of this perceived reality help to shape our mental models, in turn affecting the decision rules by which we operate, thereby influencing decisions themselves that feed back to actions in our reality. With virtual experimentation, observations and decision rules may be input as data and assumptions into our models, and our output may inform policies for making decisions. But the most critical aspect of virtual experimentation is learning from the reflexive relationship between such experimentation and our own mental models. The diagram in Figure 2 is adapted from Sterman (2000).

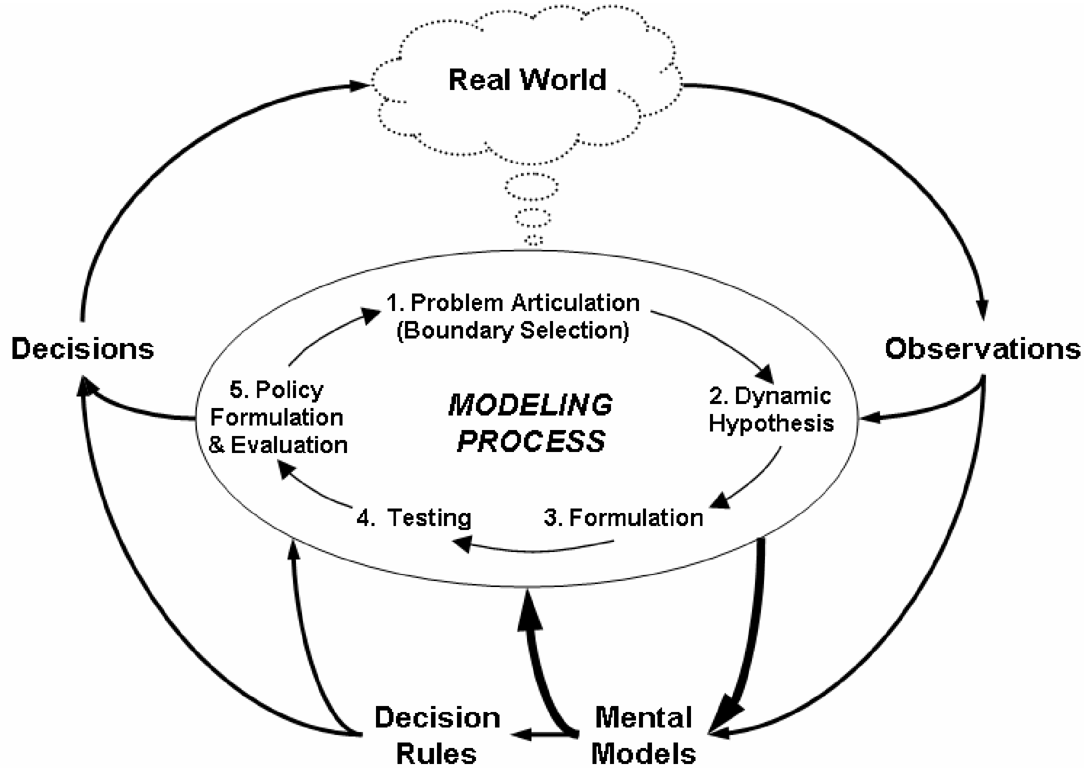

Figure 2. Learning from the Virtual Experiments of Modeling. [2]

This research thus outlines not only the development of abstract models, but also their calibration to qualitative and quantitative observations for Danville, Illinois. In addition to the in-depth interviews, I spent a significant amount of time engaged in participant observation and conversations with residents living in and around Danville. In the context of virtual experimentation (Figure 2), such real world observations served to guide model development through the adjustment of my own mental models.

[2] Adapted from Sterman (2000).

CHAPTER 2. THEORY

As introduced in the previous chapter, the problem of spatial disparity is not new. Indeed, in the taxonomy of Rittel and Webber (1973), these urban inequities constitute a “wicked” problem that defies definitive formulation, let alone a definitive answer. In contrast to “benign” problems such as logistics and infrastructure concerns, wicked problems are not readily addressed with formal modeling tools. The formulation of a wicked problem biases the solutions that may emerge. Indeed, inquiry into such problems may produce a refined and revisited problem formulation, rather than a solution per se.

A variety of classic ethnographic inquiries have demonstrated the persistence and wickedness of the problem of spatial disparity (Hollingshead 1949, Rainwater 1970, and Ley 1974). Hollingshead (1949) focused on how class and kinship ties exacerbated stratification in a Midwestern community, recognizing the intergenerational perpetuation of social economic inequities. Rainwater (1970) set out to report on the Pruitt-Igoe project in St. Louis, but emerged with still deeper questions about the embeddedness of the urban underclass. Ley (1974) also experienced a shift in understanding of the problem as he spent more time with an inner-city neighborhood in Philadelphia. For Ley (1974), this shift went from seeing the community’s role in perpetuating stereotypes, to the role of individual defensive behaviors under uncertainty. This emphasis on the individual contrasts sharply with interpretations of urban inequities as largely resulting from macroeconomic structural forces (e.g., Wilson 1987).

While not recognized by Rittel and Webber (1973), modeling methods that recognize the autonomy, heterogeneity, and stochasticity of individual actions can assist the process of iterative inquiry into wicked problems. Moreover, such methods can also bridge the gap between macro and micro interpretations of persistent social dilemmas (Schelling 1978). As described in the next secion, the use of agent-based models for virtual experimentation provides a rigorous engine for accelerating the cycle of inquiry into the wicked problem of spatial disparity.

2.1 Agent-Based Modeling

Agent based models (ABMs) may be particularly helpful in developing sense of place theory through operational models that represent individual, subjective behavior as well as aggregate, objective aspects of sense of place. Indeed, Henrickson and McKelvey (2002) advocate the use of ABMs to develop social theory congruent with new understandings of complexity science as highly contextual. Several enthusiastic authors have helped to situate agent-based modeling at the forefront of a new social science based upon complex system dynamics (Axelrod 1997, Holland 1998, and Schweitzer 2003). The term "agent" is only a minor modification of "individual-based modeling" as more widely practiced and established in the field of ecology (Grimm and Railsback 2005).

An agent is a software object that operates independently and interactively with the rest of the simulated system. The capabilities of a software agent range from simple stimulus-response behavior to more complex evaluation of potential actions. An agent may be mobile or stationary, with uniquely defined attributes and behaviors. Figure 3 illustrates relationships and structures in an agent-based modeling framework.

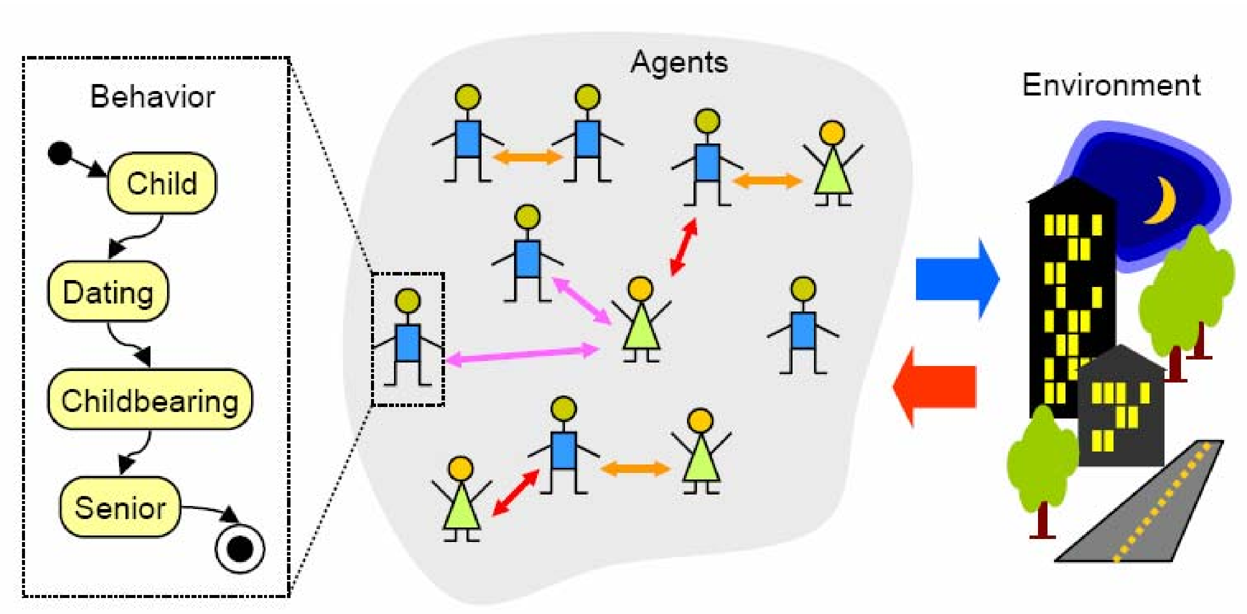

Figure 3. Agent-Based Modeling Framework. [3]

The above visualization reveals that each agent may contain structure (as indicated by the "statechart" of life stages at left), attributes and rules governing its behavior. The center and images reveals how an agent may interact with other agents as well as with the surrounding environment, respectively. A visual overview of the use of agent-based models in the java-based AnyLogic simulation environment is provided by Borshchev and Popkov (2006).

[3] This depiction of the agent-based modeling framework was developed for visualization purposes by AnyLogic. http://www.xjtek.com

2.1.1 An Emerging Methodology

Agent-based modeling has emerged in a major way with the advent of object-oriented programming techniques that enable individual agents or decision-making units to operate independently and interact with each other. Agents utilize rules for decision-making, they may have memory and may be mobile and interact in non-Euclidean space such as along social or structural networks. The flexibility of the agent-based framework has triggered a broad interest in the social sciences. While cybernetics and artificial intelligence allowed us to think about abstract representations of the mind, agent-based frameworks allow us to think about representing entire societies (e.g., Epstein and Axtell 1996). The computational load of spatially explicit agent based modeling is generally higher than other modeling approaches due to the "massively parallel microworlds" (to borrow from Resnick 1994) of spatial units or individuals.

The advantages of agent-based approaches lie in their flexible framework, ability to accommodate a variety of scales, and to capture a wide array of diverse individual behavior. Such heterogeneity and flexibility appeals more broadly to the humanistic side of social sciences that has historically been reluctant to embrace highly aggregate and generalized statistical or analytical approaches. Moreover, the flexibility does allow for a variety of theory development – theories do not need to be based on equations; rather, they can incorporate individual everyday heuristics and examine how such individual behavior produces aggregate patterns.

Critics of agent-based modeling emphasize the computational and conceptual challenges of trying to represent too much detail. The choice of the level of detail, model scope, decision-making parameters, etc., has always been a challenge for modelers. Opening the door to greater flexibility puts greater onus on the part of the modeler to fully think through the problem at hand and abstract accordingly. Because agent-based approaches are still in development, limited guidance is available in the form of established methodologies for model building and verification (Grimm and Railsback 2005, Grimm et al. 2005).

Agent-based modeling differs significantly from a continuous differential equation-based modeling. It is common in the field of system dynamics (Sterman 2000) to model a problem as a set of differential equations, represented as stocks of accumulation (boxes) and flows of change over time (arrows). Excellent introductions to system dynamics are available that also demonstrate the user-friendly Vensim (Sterman 2000) and STELLA (Hannon and Ruth 2001) icon-based software. In advocating an integral-based system representation, the father of system dynamics has observed that "nature only integrates" (Forrester 1996).[4]

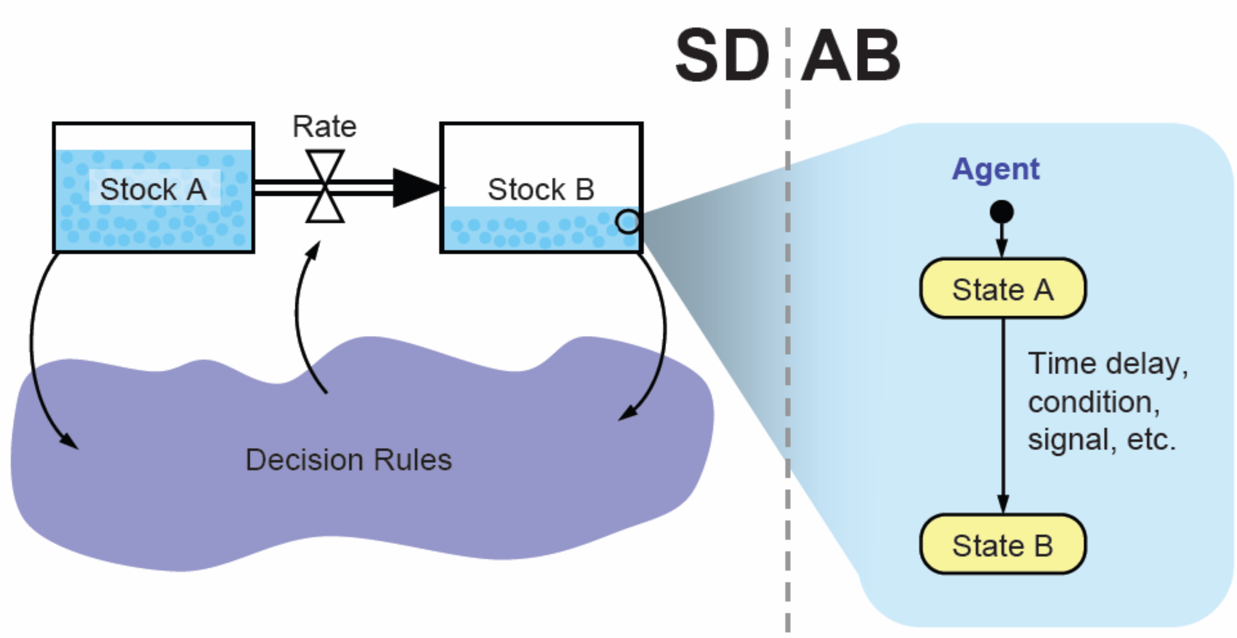

Figure 4. Conceptualizing Agents as Part of Stocks and Flows[5]

In contrast to the stock and flow representation, the agent-based (AB) framework presents a view of the problem universe as a hierarchy of decision rules. Figure 4 illustrates a system dynamics (SD) conception of problems as stocks and flows guided by decision rules using information feedback. The agent may be conceived as a discrete element within the aggregate stock. Therefore, instead of representing decision rules at an aggregate level, agent behaviors may be represented as a sequence of internal states represented by the flow diagram at right in Figure 4. In this way, agent-based modeling is essentially algorithm based modeling, where an algorithm refers to a set of instructions (a.k.a. rules, heuristics) guiding agent behavior.

---

[4] The more complete context for Forrester (1996, p. 27)'s assertion is as follows: "One might ask how it is possible to teach behavior of complex dynamic systems in K-12 when the subject has usually been reserved for college and graduate schools. The answer lies in having realized that the mathematics of differential equations has been standing in the way." ... "Differential equations are difficult, weak, confusing, and unrealistic. They often mislead students as to the nature of systems. Mathematicians have had difficulty defining a derivative and there is a reason. Derivatives do not exist except in a mathematician's imagination. Nowhere in nature does nature take a derivative. Nature only integrates, that is, accumulates. Casting behavior in terms of differential equations leaves many students with an ambiguous or even reversed sense of the direction of causality. I have had MIT students argue that water flows out of the faucet because the level of water in the glass is rising; that seems natural to them if the flow has been defined as the derivative of the water level in the glass."
[5] This diagram was presented in Borshchev and Popkov (2006) from AnyLogic (http://www.xjtek.com).

For the purpose of understanding the social influence on neighborhood choice, an agent-based framework is practically imperative. While certain aspects of social processes could be represented with alternative, aggregate methods, the ability to model social interactions in a spatial context necessitates the object-oriented agent-based framework.

Models are abstractions of reality, regardless of how intricately represented. Much insight may therefore be gleaned from parsimony, the art of leaving things out. Deciding what to put in and what to leave out is the choice that defines a modeler. The discipline of modeling encourages transparency and consistency of assumptions, but still relies on assumptions to be made. For this reason, modeling is best understood as an iterative process of formulating, testing, and refining hypotheses about system behavior.

2.1.2 Feedback and Iteration

Because wicked problems warrant an orientation to process (Rittel and Webber 1973), the modeling methodology may be considered an iterative cycle of virtual experimentation that is itself embedded in the larger context of learning and acting as social agents in the real world (see Figure 2 in the previous chapter). While the latter context is particularly relevant for considering the reflexive relationship between subject and object for a social system modeler, the modeling cycle itself enables iterative inquiry into wicked problems to build better theories of the underlying dynamics.

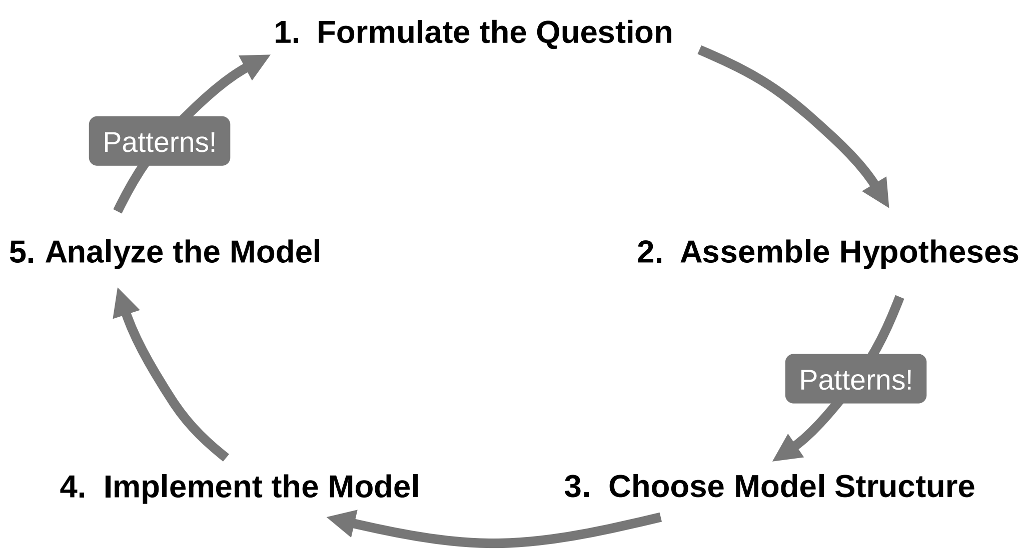

Figure 5. The Iterative Process of Pattern-Oriented Modeling. [6]

[6] Adapted from Grimm and Railsback (2005).

Figure 5 illustrates a modeling cycle (from Grimm and Railsback 2005) analogous to the cycle shown in Figure 2 (from Sterman 2000), with special emphasis on the use of patterns to guide model development as well as analysis. Such a pattern-oriented approach was adopted for this research project. According to Grimm and Railsback (2005, p.40), a pattern is "any display of order above random variation." The use of observed patterns to guide model development as well to analyze model output enables decoding of essential system information (Wiegand et al. 2003).

Grimm and Railsback (2005) have written a thorough text outlining methodology for this type of modeling, using the synonymous phrase "Individual-Based Models." In this text, they provide examples of the calibration process for agent-based models, which is still emerging. Most examples of agent-based models have been theoretical, pedagogical "toy models" (e.g., Resnick 1994, Epstein and Axtell 1996) illustrating how collective system structures emerge from decentralized individual behaviors (Axelrod 1997, Holland 1998).

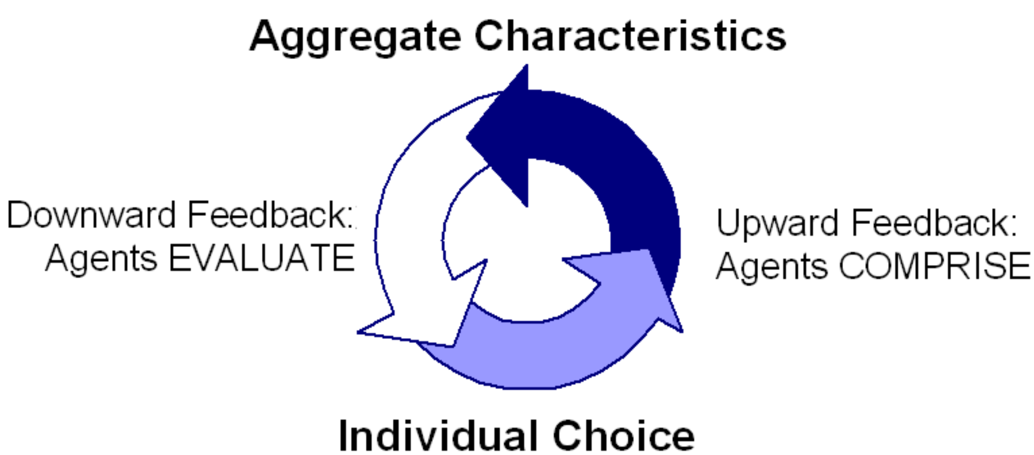

Figure 6. Feedback Across Scale in Agent-Based Models

As visualized in Figure 6, the dynamics of agent-based models involve implicit feedback across the scale of the system. Individual choices of the agents provide the dynamics of the system. This generates upward feedback, in the sense that agents *comprise* larger populations at various scales of aggregation. These aggregate characteristics or *states* change based upon the decisions made at the individual level. Such state changes at the aggregate then feed back to influence other choices, as agents evaluate their environment based upon such states – for example, neighborhood affluence.

Agent-based approaches have been used to transcend scalar difficulties in geography (Batty 2005), test effects of heterogeneity (Brown and Robinson 2006), and simulate mobile objects in dynamic landscapes (Westervelt and Hopkins 1999). Parker et al. (2002) offer criteria that may be utilized to determine if an agent-based model is spatially explicit or warrants spatially explicit representation. Such criteria include: dependence of results on relocation of objects under study, location-specific representation or formulation, and shifts in spatial outcomes over time. The benefits of spatially explicit modeling depend primarily on the problem at hand, as they tend to increase the computational requirements.[7] As agent heterogeneity shifts modeling emphasis from the aggregate to individual-level knowledge, spatial heterogeneity plays a similar role in refocusing understanding of a complex system through spatially explicit representation. An abstract notion of space is utilized in many agent-based models, and further efforts are underway to integrate agent-based models with GIS (Gimblett 2002).

2.2 Social Networks

This section outlines the basis for the study of social networks as they relate to the question of neighborhood choice. While much research in these areas has taken place within geography, sociology, and other relevant disciplines, the scope of this project offers a means of addressing disconnects between research domains – connecting qualitative and quantitative insights, spatially extending social networks, embedding individual choice within those networks, and evolving social networks over time as they influence neighborhood choice.

As introduced earlier, a social network refers to a set of interpersonal relationships. In this project, the relationships are defined as communication links that evolve in the face of neighborhood and community migration. Although networks are powerful, they are difficult to define and measure, let alone simulate over time. Research on social networks has tended to be ethnographic (Rowe and Wolch 1990, Gilbert 1998), empirical as in much of social network analysis (Wasserman and Faust 1994), or abstract as for emerging simulation techniques (Watts

[7] To address computational requirements, software tools such as StarLogo, SWARM, and RePast have enabled efficient simulation of a wide variety of spatially explicit agent-based models.

and Strogatz 1998, Barabasi and Albert 1999). Emerging indirect estimation techniques (Conley and Topa 2003) enable calibration of abstract models from spatial socioeconomic data.

Studies of social network structure ahve demonstrated the significance of homophily, or self-sorting according to similarity. Homophily has been demonstrated by similarity along race, age, religion, education, occupation, and gender lines, with geographic proximity and family ties creating opportunities for such self-sorting connections to form (McPherson et al. 2001). Lazar et al. (2002) note the importance of being able to detect others' type that enables homophily in the first place. In a related study, Macy et al. (2003) examine how polarization or segregated clustering can occur in the absence of resource competition. They encode an attractor network in their model, such that agents are attracted to others of similar states and are also influenced by others. While a novel approach, their work is consistent with prior findings of polarization under the principles of structural balance.

The research project described herein builds on prior studies of social network structure. The study of social networks is rich with methods of structural analysis that are based upon graph theory as stimulated in large part by Erdos and Renyi (1960). Assumptions about network connectivity may also be derived from the General Social Survey (Davis et al. 2005), which is available longitudinally from 1972 to 2004. Conley and Topa (2002, 2003) utilize data from the General Social Survey to model the degree of connectivity among individuals.

Social networks enable a purely relational perspective, where one of the biggest challenges is simply identifying relationships to be analyzed. For this project, the relationship is one of communication. Two areas of emerging research involve *spatial* and *dynamic* social networks. Next I outline the direction of this research and how the two foci may be combined.

2.2.1 Social Networks in Geography

During the 1960s, Stanley Milgram (1967) conducted an experiment that came to undergird "small world" theory. Milgram gave research participants in Kansas and Nebraska a letter describing a target person in Massachusetts. If the participant knew the target on a personal basis, he/she was asked to send the letter directly to that person. Otherwise, the participants were to give the letter to a personal acquaintance who was more likely to know the person. Of the letters that were returned, the median number of intermediate links was 5.5, rounding up to the cliche "six degrees of separation."

What is striking about Milgram's research is that it was inherently geographical in nature, intended to measure social distance spanning geographic locations between arbitrarily chosen individuals. And yet the subsequent research in social networks has tended to involve purely "relational" space, without consideration of geographic distance. Meanwhile, geographers have pursued the study of social networks and other networks in both empirical and ethnographic forms. While complete social networks are difficult both to define and to contain, the concept of networks has proved useful for understanding the role that individual networks play in shaping life opportunities (Rowe and Wolch 1990, Gilbert 1998, Peake 1995).

Adams (1995, 1998) brings the concept of human extensibility to light, emphasizing the myriad of connections across virtual space and time as well as in real or observable space and time. The concept of extensibility sheds light on the difficulty of identifying what it is that constitutes a link, and what it is that constitutes space in the geographies of our everyday social experience. Since a connection in social space is enabled by human-human interaction, simulation techniques at the individual or agent-based level offer a means to grapple with the inherent complexity of our many linkages.

Before the semantics of social networks became commonplace, mathematician Ron Atkin (1974a, 1974b) introduced a technique called Q-analysis that was then taken up by geographers Jeff Johnson (1981) and Peter Gould (1981). Q-analysis involves assessing structural relationships in multi-dimensional spaces that transcend Euclidean space. Atkin (1974a, 1974b) proposes that such representation could provide a more holistic view of urban areas, indeed one that enables simultaneous consideration of both global and local properties. Further, he argues that such structural relationships correspond to the intuitive experience of the urban environment.

The structural work of Laumann and others (Laumann 1973, Laumann and Pappi 1973, and Laumann et al. 1977) provides foundations for an understanding of the social influence structures at play in urban communities. Indeed, Laumann (1973) argues that a sound understanding of structure must precede an examination of social change. While much more structural research is warranted (and certainly geographers are interested in *spatial* structure), it is the behavior that comes from structure and that ultimately changes structure that inspires research into the *dynamics* of social networks.

2.2.2 Dynamic Social Networks

The dynamics of social networks may be considered in two ways – the dynamics of behavior within a network structure, or the evolution of the network itself over time. Degenne and Forse (1999) introduce dynamics of social networks in terms of diffusion. While providing useful frameworks, they note the need for much more work in the area of dynamics. Burt (1987) explores the implications for dynamics of diffusion from cohesion versus structural equivalence approaches to understanding networks. To connect structure with dynamics, White (2004) presents a synthesis of social network theory in relation to social dynamics, including a detailed conception of how statics and dynamics operate in balance through many patterns witnessed in social network theory. In a related work, White et al. (2004) focus on cohesive network topologies in both organizational contexts and emergent fields, and the ways in which these interact.

Applying an epidemiological notion of diffusion to the social context, Granovetter (1978) demonstrates the utility of threshold models in understanding collective behavior. Social thresholds refer to the minimum fraction of one's peers who have made a decision before the individual in question does. In a similar manner, Crane (1991) utilizes a contagion model to examine the nonlinear social effects of neighborhood dynamics as behaviors transmit among members.

The dynamics of social networks have been examined from a number of perspectives. Epstein and Axtell (1996) introduce ways in which simulated agents can represent human connections and interactions over time. Indeed, they emphasize the importance of transients and dynamics more than the quest for equilibrium conditions. With a similar approach to modeling, Young (1998) focuses on theory underlying such individual-based conceptions and the institutions that result from such interaction. The extent to which individual preferences are selfish or altruistic has also been examined in recent simulation studies (Bowles 2001, Lazar et al. 2002) of the simultaneity and reflexivity of network evolution and individual preference evolution. These studies reveal not only that trends toward conformity within groups can sustain difference between groups (see also Young 2001), but also that within-group cohesion may result from socially influenced altruism that runs counter to selfish motives.

Watts (1999a, 1999b) and Watts and Strogatz (1998) explore self-organizing dynamics of network formation, emphasizing that where self-organization occurs, the resulting structure lies between randomness and order. The prime example of such a mix of randomness and order is the small world network inspired by Milgram (1967), in which local clusters are dense but are connected globally through a few cross-cutting links between hubs. Accordingly, Watts (1999a) developed algorithms for evolving small worlds in which interpersonal connections are locally dense (e.g., most of my friends are also friends) but globally sparse (e.g., everybody is not directly connected to everybody else). This small world lies in an interesting region between complete subgroup isolation and complete network connectivity. While the suitability of the small world structure to describe real-world social networks is still under evaluation, it is one of the most promising quantifiable theories of social structure. A critical element of Watts' (1999a, 1999b) small world formulation is a probability of "rewiring" social connections from an initially ordered structure (usually a ring lattice of one-dimensional connectivity).

A parallel development published shortly after the Watts and Strogatz' (1998) expose of small world network dynamics was the emergent and pervasive scale free network structure identified by Barabasi and Albert (1999). The scale free network exhibits a power-law degree distribution such that a few nodes have very high connectivity and most nodes have very low connectivity. The fascinating part of this structure was the dynamics of preferential attachment that enabled hubs to emerge. Expanding on the relevance of these network dynamics, Barabasi (2002) and Buchanan (2002) recently helped to popularize the new "science of networks" ranging from the curious small-world phenomenon to the apparent prevalence of scale-free networks in all kinds of systems.

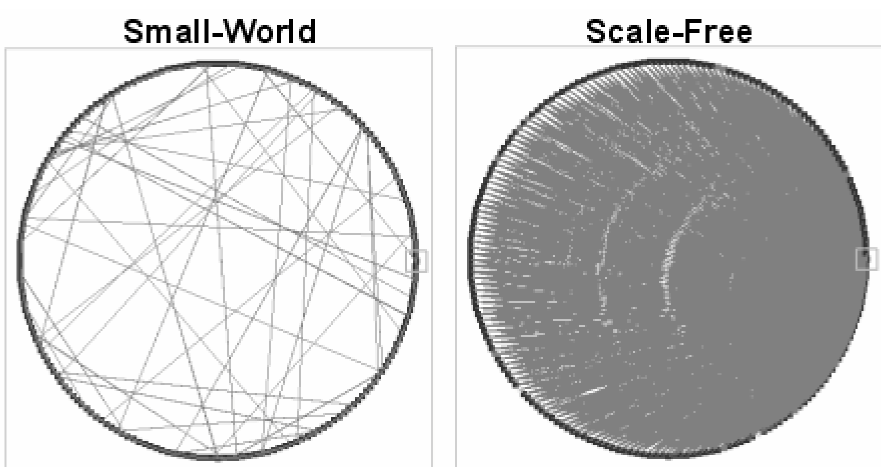

Figure 7. Example Small-World and Scale-Free Networks

As illustrated in Figure 7, a small world network (left) is primarily local with a few long-range connections (Watts 1999a). A scale-free network (right) involves preferential attachment so that popular people become even more popular (Barabasi and Albert 1999). Small worlds may be scale free and vice versa.

As it turns out, small worlds and scale free networks are not antagonistic, such that many networks may exhibit both characteristics. The small world is defined by a short average distance (e.g., number of handshakes) to reach anybody else, while remaining locally dense and globally sparse. The scale free network is defined by a degree distribution that results from preferential attachment to a few popular hubs. It may be that these hubs can also provide the criteria for small-world phenomena. Figure 7 above illustrates a comparison of the two networks for the same average number of connections, showing how the network on the left exhibits a few randomly placed long-range connections, while the scale free is dominated by a central hub at the right of the graph. Thus in this case the two networks are not equivalent.

The study of network dynamics adds complexity to social network analysis, a field that is already full of techniques to test social relationships. And yet, as evidenced by the above literature, it offers the prospect of additional insight in understanding the simultaneity of how the network influences the individual, and how the individual influences the network. By exploring relationships at the level of individual interaction, we can learn about behavior over time at both micro and macro social network scales of analysis.

2.3 Conceptual Integration

From a simulation perspective, much has been done to examine the emergence of social norms from individual interactions (Durlauf and Young 2001). Thus far such simulations have tended to be highly abstract. Moreover, they have largely been disconnected from a consideration of physical space. With fieldwork as well as simulation techniques, this research serves to integrate these pieces to address the problem of community fragmentation.

2.3.1 Learning is Social

A variety of theories support the agent-based framework as introduced in social science. Latane's (1981) social impact theory holds that social influence is a function of strength, immediacy, and the number of influencing agents. The dynamic extension of this theory (Latane

1996) has helped inspire interesting simulations of opinion dynamics (Weisbuch et al. 2001), exploring how individual decision rules for learning and adapting under social influence can create emergent social norms. As an alternative approach to neoclassical economics, the study of individual choices in an adaptive and adapting environment enables a foundation for a new social economics (Durlauf and Young 2001). As some researchers argue that all learning is social (Kennedy and Eberhart 2001), simulated experiments not only shed light on emergent structures but also on evolution of subjective perspective. Further, Kenrick et al. (2003) outline a theory of dynamic evolutionary psychology that emphasizes both the social context of decision rules and their evolution, as well as attributes of the individual and the problem.

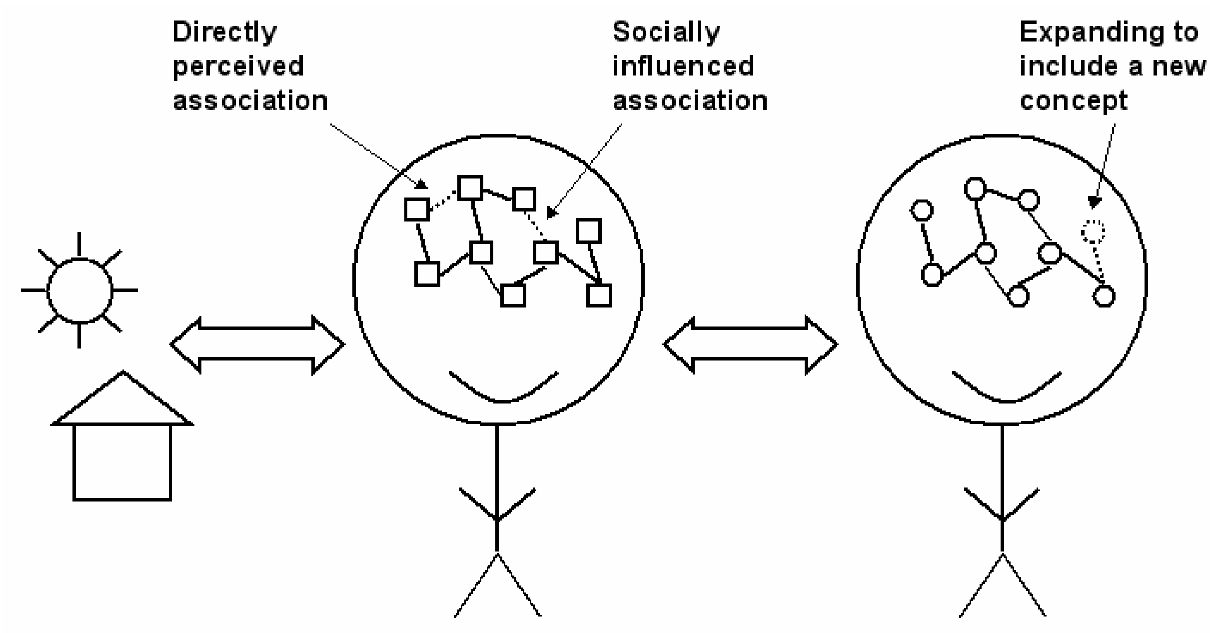

Figure 8. Visualization of Interactive Learning. [8]

The visualization in Figure 8 demonstrates how individual perceptions may be updated through new concept associations, either directly perceived or socially communicated. The adjustment of individual perceptions could be represented as a memory effect that decays over time. This decay is analogous to the notion of time discounting in economic systems, where a dollar is worth more today than tomorrow. There is no time like the present, and as we get further away from experiences of our past, we may effectively discount it as we do the future, especially if it is not reinforced by current behavior. Thus a corollary of social learning is

[8] Adapted from A. Hubler lecture. April 2004. PHYS 521: Advanced Nonlinear Dynamics. <http://www.how-why.com>

selective forgetting, which might be considered social discounting. Such discounting relates back to an individual's sense of place. Indeed, Hannon (1987, 1994) demonstrates that such geographic discounting of attention or concern is not restricted to humans, but also applies to other species. To address social learning and discounting in a spatial context requires an explicit consideration of social networks in a spatial context.

2.3.2 Social Networks are Spatial

Spatial simulations of neighborhood networks date back to the cellular automata simulations of Hagerstrand (1965) and Schelling (1971, 1978). Cellular automata are discrete cells located on a grid that update their state based on their previous state and the state of their neighbors (Shalizi 2005). Cellular automata are a sort of precursor to the more richly structured decision rules of mobile interactive agents. Hagerstrand (1965) utilized empirical data on telephone network density to explore stochastic simulations of spatial diffusion of farming subsidies in Sweden. Schelling (1971, 1978) employed an abstract cellular framework for examining the emergence of segregation from low thresholds of preference for similar neighbors. While such simulation utilizes abstract space in the form of a uniform grid of household locations, it enables the development of intuition about the conditions under which spatial clustering emerge and the degree of contingency in its patterns.

In a related abstract spatial simulation, Arthur (1988) explored the uniqueness that can play out as a result of historical path dependence from industrial location decisions. While Arthur (1988) illustrated the potency of a simplified abstract model, geographers with expertise in Geographic Information Systems (GIS) are ready to add realism to spatial simulations. More recently, Dibble and Feldman (2004) presented a computational laboratory in which simulated agents may interact in networks across abstract or empirical space using GIS templates.

The argument that social networks are spatial is at once obvious and yet superficial to sociologists who have resisted it with a pure focus on relational space instead. And as geographers recognize, life space is far from Euclidean. Yet abstractions of space in just two dimensions can prove pragmatic for navigating our world, much as abstractions of relational space in social networks are relevant to human understanding. The combination of these spaces offers a way to make the small world (Watts 1999a) rewiring probability endogenous in a spatially explicit environment that incorporates individual choices about whether to leave neighborhoods. A major contribution of the research described herein to the field of social network analysis is this dynamic evolution of networks in a spatially explicit environment.

2.3.3 Neighborhood Choice

Although embedding individual perceptions in agents who are in turn embedded in social space may be a rich representation, to put a simulated agent in action requires behavioral rules connecting attitude to choice. A variety of rules have been constructed based on rational expectations and utility theory, or on the simple assumption of random exchanges. In contrast to rational rule structures, alternative or complementary approaches such as case-based reasoning draw upon memory and association through analogy. The case-based reasoning algorithm requires a metric of similarity by which to measure new experiences. Agents are, in essence, reminded of past experiences when they encounter similar situations (Schank 1999). This phenomenon extends the notion of discounting to the level of individual choice, such that associations not reinforced with new experiences are forgotten over time.

Behavioral and cognitive geographers interested in the connection between cognition and choice have considered personal construct theory, which holds that personal constructs are used to structure individual experiences of life events (Preston and Taylor 1981, Moore and Golledge 1976). Distinct variations emerge therefore according to "family life cycle," as different constructs emerge with different experience sets (Preston and Taylor 1981). Further, Desbarats (1983) notes the need to incorporate constraints on choice that lie outside of cognition, such as situational and institutional constraints. Couclelis (1986) examines alternative models of spatial decision and behavior within a common theoretical framework, highlighting that useful insights may be gleaned from comparing such models. The discrete choice methodology outlined by Ben-Akiva and Lerman (1985) and McFadden (1991) has been widely applied to urban and transportation problems. Applications have demonstrated the use of logit models for discrete choice selection of social connections (Van de Bunt et al. 1999) and for choices made in the urban context (Waddell et al. 2003, Paez and Scott 2007).

This project has not endeavored to invent a new theory of choice. Rather, the goal was simply to test the influence of social networks on neighborhood choice. Apart from external influences, the decision to leave a neighborhood is a choice influenced by perceptions that are mediated in the social context. Emphasizing that neighborhood stability tends to be highly valued, Aitken (1990) explored perceptions and evaluations of neighborhood change. To the extent that such evaluations lead to tangible moving decisions, a further implication is the reinforcing impact of neighborhood out-migration on social network structure. In addition to direct cues of out-migration, such as “For Sale” signs and vacant homes, such departure ripples through the fabric of the community. This is not to say that once an individual household leaves a neighborhood, they are automatically “deleted” from their social network. But the tie may weaken relative to social connections that derive in part from household proximity. Moreover, if households leave a community altogether, the fragmentation effect on social network structure would be even stronger.

CHAPTER 3. EMPIRICAL CONTEXT

This research was motivated by the question of why people decide to *leave* certain neighborhoods in a community. How was it that some places, initially desirable, became undesirable over time? What role did social influence have in shaping individual choices of where to live? While a simulation model would be useful for addressing this question, a real-world case community was needed to "exercise" the model, reveal its limitations, and ground its implications.

3.1 Study Site

Located two hours south of Chicago in Illinois, Danville is a self-contained community of over 30,000 people. It occupies the seat of Vermilion county, whose eastern boundary is shared with Indiana.

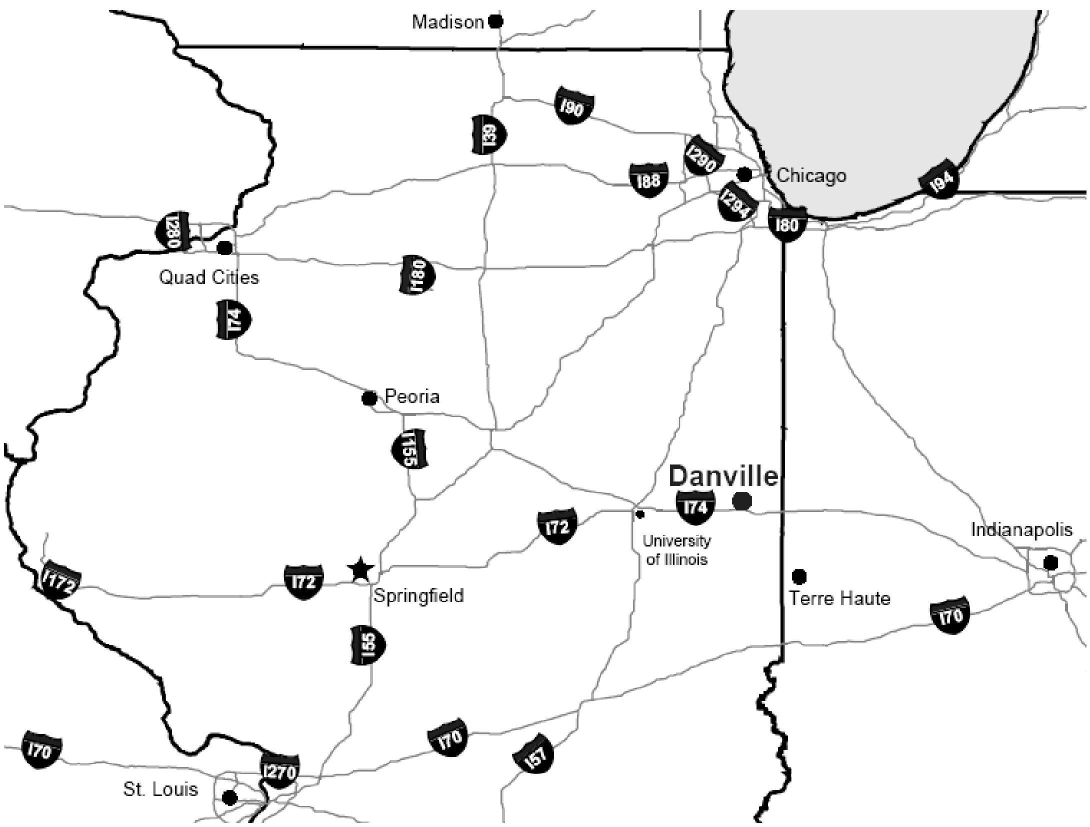

Figure 9. Relative Location of Study Site: Danville, Illinois, USA. [9]

[9] The map in Figure 9 was obtained from City of Danville: <http://www.cityofdanville.org>

3.1.1 Recent Trends

Over the past few decades, Danville has experienced population decline and stagnation, but still experiences sprawl in the form of new fringe development. The trends in population and land area over time in Figure 10 reveal that Danville's land area has nearly doubled in the last half-century, though its population is the same as it was in 1920. Danville's population peaked in 1970, with well over 40,000 people. As one city official put it, "we're a town of 30,000 with an infrastructure for 60,000." Danville is still the largest community in Vermilion county, with over 40% of the county's population.

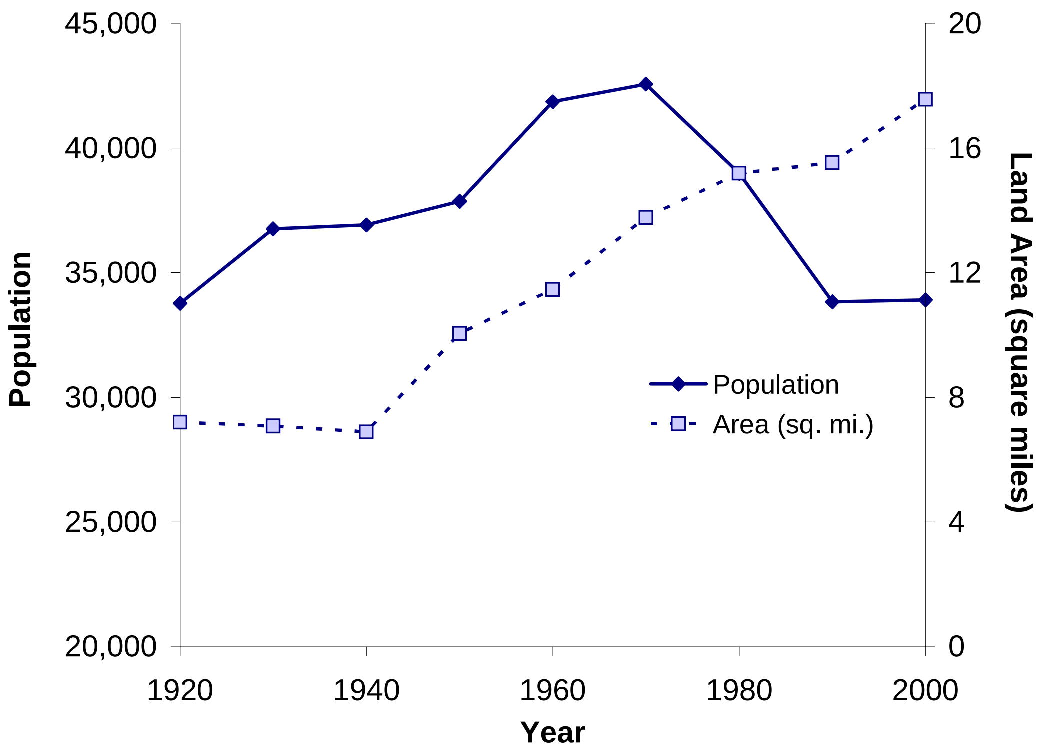

Figure 10. Trends in Population and Land Area for Danville, IL, USA

Until the 1980s Danville was a strong working class town hosting industries such as General Motors, General Electric, and ALCOA. The exodus of these industries in search of lower wages echoed the classic *Roger and Me* story told by Michael Moore (1989), leaving masses of people unemployed. This created a tension between attachment to hometown roots and the struggle to find work. Many middle-class workers left to find jobs elsewhere.

Today the city of Danville is struggling to transform itself to remain competitive in a global economy. Although much of Danville's core population lives in poverty, it maintains a strong class of affluent professionals and doctors who serve the community's hospitals and retirement homes. In recent years, Danville's prison has extended its social network to residents of nearby cities such as Chicago. With recent gentrification efforts, many low-income residents of Chicago's housing projects have chosen to relocate to Danville for subsidized Section 8 housing. Such in-migration keeps Danville's population from declining further, but raises concern for long-time Danville residents as social services become taxed and unfamiliar faces appear in neighborhoods and schools. To stabilize the neighborhood dynamics, the mayor and the city of Danville are actively promoting neighborhood revitalization through resident-organized neighborhood associations.

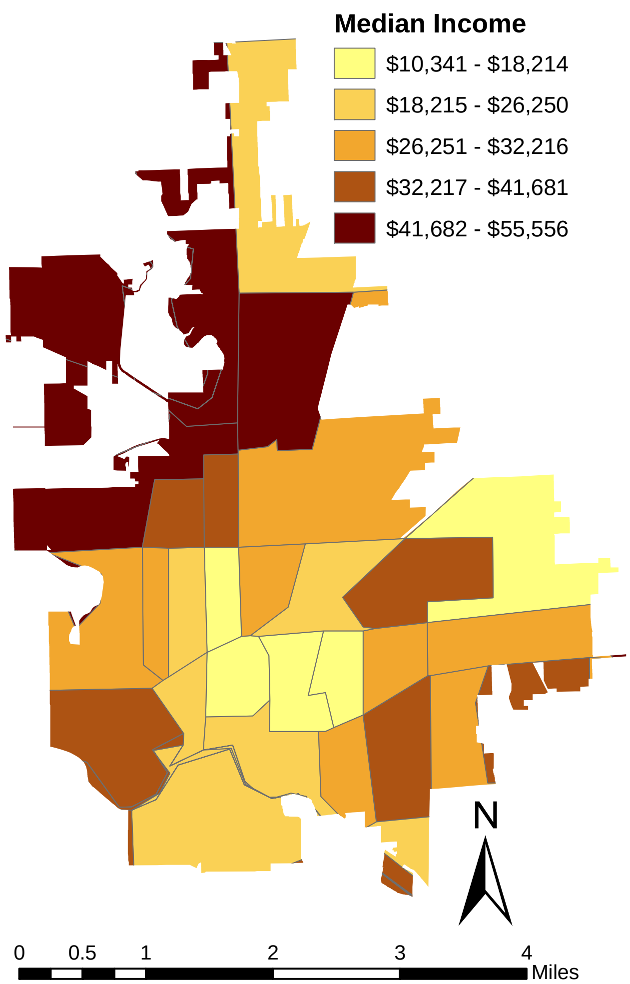

Figure 11. Danville Median Income by BlockGroup

The blockgroups outlined in Figure 11 were visually clipped to the map of Danville boundaries exhibited in Figure 12 below, excluding the Easternmost area. This Easternmost area was excluded from model development (see section 3.3.2 below) because it is not a significant residential area and it creates a spatial discontinuity. Another spatial discontinuity is in the Northwest area across from a man-made lake (a modern "moat" of sorts), but this area was retained due to its residential significance.

Danville has a very clear sense of the spatial division between the haves and have-nots. The North and Northwest areas are affluent, while the South and East areas of town are modest to very poor. This distinction alongside continued growth in land area made Danville a solid case study for simulating the social dynamics of spatial disparity through neighborhood network evolution.

3.1.2 Recent Studies

A recent report on the workforce in Vermilion County outlined three scenarios for the near-term (circa 2010) future (Judy and Lommel 2002, pp. 43-45). The first scenario, titled "no surprises" highlights the sobering trend of declining employment in the face of global competition, and a steady flow of commuters who live in Vermilion county but work in the neighboring Champaign county. The second scenario envisioned the realization of hopeful economic development plans, capitalizing on demand for health care and niche industrial opportunities to reverse the "brain drain" of human capital that has taken its toll in the past three decades. The third scenario considered what could happen if "too many things go wrong", outlining vicious cycles of reinforcing feedback that could further deteriorate the educational system, employment prospects, and accelerate the "brain drain" of young and experienced residents. While the authors (Judy and Lommel 2002) noted that "every effort should be made to avoid" the third scenario of rapid decline, it is an unfortunately plausible scenario for the community. Between 1990 and 2000, Vermilion county lost over 4000 residents, nearly 5% of its population.

In 2005, the city of Danville conducted a survey of the community, eliciting responses from concerned citizens.[10] The results of this community survey include 9% (22 of 238) responses from ward 3 (see ward map in Figure 12), which encompasses the neighborhood of

[10] "City of Danville Community Surveys" Presentation dated 28 Nov 2005. Obtained from Chris Milliken.

focus for the ethnographic portion of this research. Overall, 60% of respondents had lived in their neighborhood more than 10 years. When asked what they liked best about their neighborhood, the most frequent response (34%) by far was "friendly neighbors." In ward 3, this was listed even more frequently (by 9 of 22 respondents, 41%). In contrast, the responses to the question of what was least liked in their neighborhood revealed a variety of responses, with different priorities for different areas. For example, rental housing was cited in 8% of overall responses, but was the top concern listed for ward 3 (by 7 of 22 respondents, 32%).

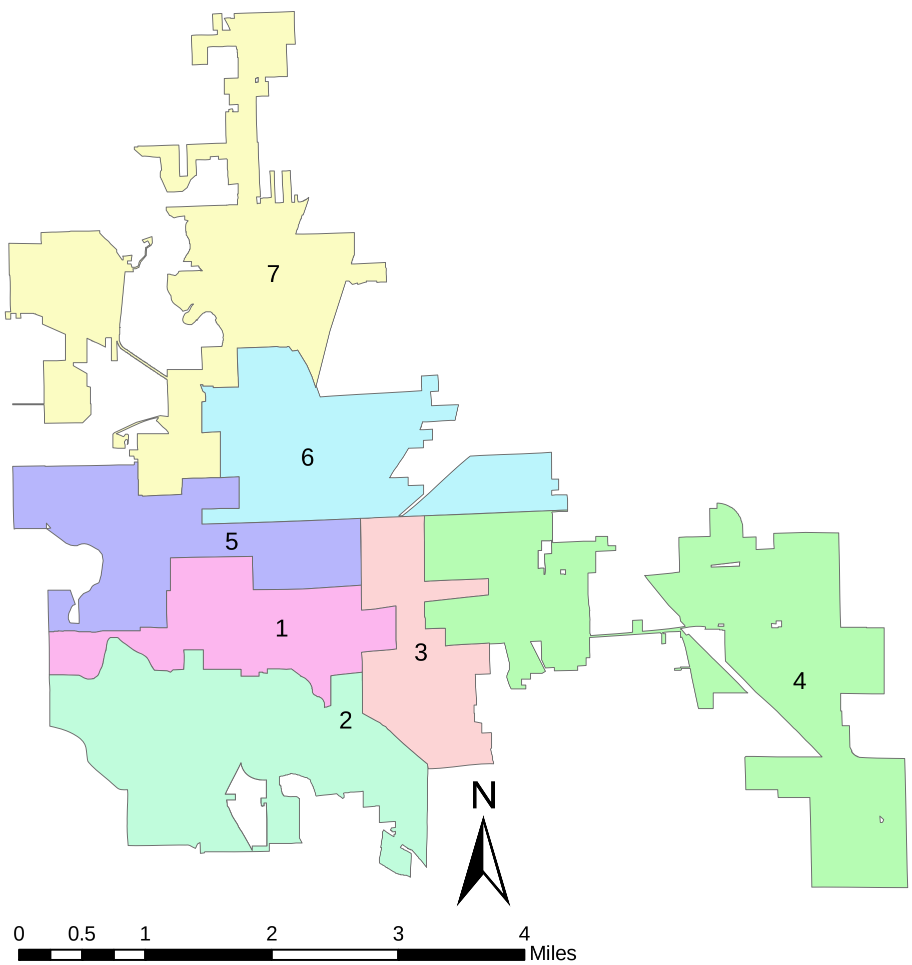

Figure 12. Map of Danville Wards

3.2 Data Sources

Table 2 provides an overview of the Danville data that were used in this study. Qualitative data in the form of interviews and participant observation were collected explicitly for this project, and quantitative data were obtained from the City of Danville and the Census.

Table 2. Qualitative and Quantitative Data Collected for Danville

	Qualitative Data		Quantitative Data		
Form:	Interview	Observation	Owner List	Parcel GIS	Area GIS
Date:	2005	2005	2001, ‘03, ‘05	2005	2000
Source:	Residents	Visits	City of Danville		Census

Direct quotes provided data as well as survey responses (e.g., to the Census). Moreover, the direct experience itself (as generated from participant-observation) biased the researcher’s empathy for the subject of research. This subjective intuition then better informed the individual-based heuristics, because they were encoded from the modeler’s perception of the relevant behaviors. Indeed, social scientists are increasingly recognizing the role of ethnography in creating more robust models (Agar 2004).

3.2.1 Qualitative Data

The ethnographic fieldwork (interviews and participant observation) took place during the summer and fall of 2005. Initial activities involved identification of research participants appropriate for in-depth interviews. This process began with informal interviews with city officials and representatives from Vermilion Advantage, an institution to promote economic development in Vermilion County. City officials and community advocates were consulted for preliminary identification of an appropriate target neighborhood in flux. Then a specific neighborhood undergoing substantial change was identified as a focus for interviews. This neighborhood had recently formed a neighborhood association to identify and act on local needs. The interview protocol in Appendix A provides an overview of the open-ended questions that were used. Interview length varied between 45 and 90 minutes per session. Each session was recorded and transcribed. Due to the in-depth nature of the interviews and the availability of complementary data (via participant observation and city GIS datasets), eleven (two preliminary and nine formal) interviews were conducted for this research.

Resident interviews were conducted alongside processing of GIS data, during the period from September to December 2005. The qualitative interviews (see Appendix A for the interview guide) served to “ground truth” the research, reminding me of how limited the model must be relative to the scope of experience of Danville residents. The open-ended questioning provided the advantage of a smooth, natural dialogue for the interviews. However, the elicited narratives naturally diverged from the central research focus, in ways that were highly intriguing but must be explored more thoroughly in subsequent analysis. For the purpose of this research, selected experiences and statements were used to augment chief assumptions and model heuristics for decision algorithms and boundaries.

3.2.2 Quantitative Data

As outlined in Table 2 above, the City of Danville provided a variety of quantitative data used in this analysis. These data included lists of homeowner names and addresses for the years 2001, 2003, and 2005. The homeowner lists were collected in the spring of each year as part of tax assessment, and so reflected the homeowner residence for the prior calendar year. The 2001 list was used to derive homeowner locations in the year 2000, and was combined with Census data for the year 2000 to initialize the simulation model as described in the next chapter.

In addition, the city provided GIS data identifying parcel locations and type, for which the metadata are illustrated in the Appendix. Combining locations with the owner data for 2001, 2003, and 2005 enabled derivation of owner migration matrices in the intervals from 2001-2003 and 2003-2005. The methods for processing these data are described in the next section.

3.3 Data Processing

This section outlines how the data obtained from both qualitative and quantitative sources were processed to provide grounding for the simulation model. The data integration process was a significant part of this project, one that is described further in the next chapter.

3.3.1 Interview Heuristics

The neighborhood of focus was mixed in many dimensions: tenure (owners and renters), age, race, income, and family status. While research participants who provided in-depth interviews for this study spanned race, age, and family dimensions, they were all homeowners. (Income was not elicited in the interview process.) Perspectives on renters versus owners were elicited informally through participant observation of neighborhood association meetings (which were dominated by homeowners) and through conversation with city officials. The city officials openly noted that the renter-owner ratio can reach a tipping point that creates instability in terms of outmigration. Early on, therefore, the distinction between renters and owners was highlighted as one that might make a difference in the model. One participant commented that “a lot of the homes in this area are rentals, and rentals are going to attract more transient type of people.” Another noted: “The neighborhood became a lot of rental property, and I think that’s probably the biggest issues there. They’re not very selective maybe because many of them will go Section 8 housing. So they’re not taken care of in the way they might have been had they been homeowners.”

One newcomer described the neighborhood of focus as “the other side of the tracks” and emphasized that it is a humbling experience to come to terms with her own biases. Also, references to the “wrong side of town” (made by a white newcomer) corroborated the implicit spatial distinction between the “haves” and “have-nots” in Danville.

Another participant talked about the racial diversity of the neighborhood: “It’s mixed. It’s like black, white, black, white.” Despite the diversity, some racial tension was noted: “When you get right down to the nitty gritty, everybody has a prejudice and it may not show, but it can be felt. If we’re going to integrate we’re going to have to integrate right from intermarriage right on up. We’ll always have a separation until the youth, the preschoolers of today are old like I am and have outgrown all that.”

The importance of children in social networks was revealed. “Lots of kids live on this block. Teenagers, everybody was just hanging out over here like we were the spot.” “My daughter knows more people around here. Like [neighbor], she met him because she would go out there and take out the garbage and they would be out there in the garage, because my grandchildren would ride their tricycles around the garage.” Changes in neighborhood family structures were also noted: “Over the years the neighborhood has changed, and we have a younger group coming in now, that’s [got] kids and so forth. As we began to see that [older] age group leave the neighborhood, I kind of took a private stance, because I didn’t know if [the new people] were friendly or not.”

Some residents voiced concerns about neighborhood children: “When the kids were little, the were cute little things and they’d come over and talk to you when they were doing things out in the yard; but as they got older, they got a little more malicious.” “Now the little kids – well, I’m sitting out in the summertime and they are running all over and it scares me to death. I don’t say anything but I think they let them run a little too loose and it’s not that safe on [street].”

The combination of mixed racial backgrounds along with the draw of children's networks produced some tension: "I think if we would get together more often, we would know each other more. Then we would trust each other more. Because when [child] goes over [to neighbor's], I can tell they really don't trust, I'm not going to say me, but they probably concerned. I don't know if it's because of color. They pick and choose like one day she could come over then another day she can't come over. I don't know what that's about. And one day she can come over but she can't come in."

Residents spoke more broadly about what it means to be a neighbor, and to have a neighborhood. Appendix B provides additional quotes organized by theme as relevant to the neighborhood association.

"A dictionary probably says the person that lives in your area, your environment, surrounding, next door, behind you, in front of you, that's a neighbor. I consider a neighborhood as [a place] where I could come to you, you could come to me, we can help each other out instead of just going past – 'hi' – and that's it. I don't even know your name."

"I think the world has become a place that if you find something wrong with your neighborhood, the easiest thing to do is leave. The toughest thing to do is make it better. There is a great little quote: 'Adversity does not build character, it reveals it.' What I hope is over time, if we've had adversity in this community, it's not going to build character for the community but it's going to reveal the character that we already have."

These quotations were used to guide several assumptions in model development: the inclusion of race and presence of children as factors in social network choice, and the importance of a distinction between renters and owners. The general influence of social networks in stabilizing neighborhood dynamics was observed through neighborhood association activities. This ethnographic fieldwork also served to humble the modeling activity, highlighting nuances and disturbing behaviors (e.g., the influence of domestic violence on decisions to divorce and relocate) that were beyond the scope of this project. Quotes relevant to the neighborhood association were offered as outlined in Appendix B.

3.3.2 Parcel Data Processing

Although GIS data at the parcel level were readily available for the Danville context, they required processing to ensure an appropriate scope and to integrate with Census block and blockgroup boundaries. The following steps describe how the parcel GIS data were processed.

1. *Selecting Danville Parcels*. The parcel file was provided for the entire Vermilion County. From this file, parcels were selected with an address explicitly identifying the city of Danville. The resulting number of "Danville" Parcels was 16,645.

2. *Selecting Danville Place Blocks*. A shapefile was created for all housing-occupied Census blocks associated with the "place" (Census code 18563) of Danville City and projected in the same way as the parcel dataset (see Appendix C for projection). Blocks that were spatially distant from the center of Danville with few occupants were removed (namely tract 10, blockgroup 1, blocks 2,5,8,29; and tract 105, blockgroup 1, block 35). These blocks are revealed by the easternmost gray area of Figure 13, which illustrates the boundaries of the selected 786 blocks with the entire set of 991 Danville blocks.

Figure 13. Selected Danville Block Boundaries

3. *Selecting Parcels within Blocks*. A point shapefile was created for the centroid of the 16,645 Danville Parcels. This file was spatially joined to the selected Danville blocks of step 2. The resulting 13,173 successfully joined parcels were retained for the analysis. Figure 14 illustrates the residual alignment discrepancy between Parcel and Census Block datasets for a selected area of East Danville.

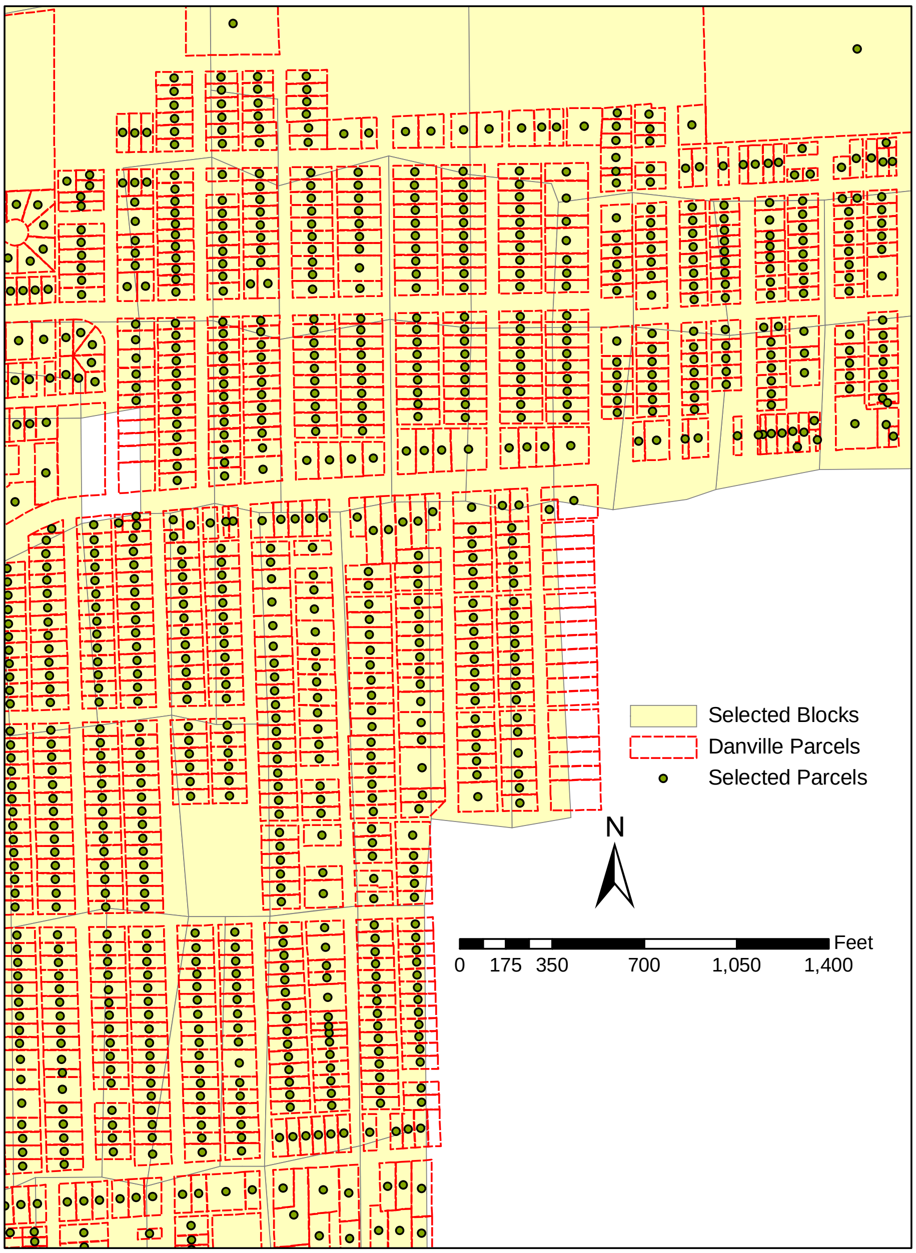

Figure 14. Difficulty Associating Parcels with Census Blocks

4. *Identifying Owner Parcels*. The 2001 list of owner-occupied addresses as provided by the city of Danville was aligned with the selected Parcel data. The resulting 7579 owner-occupied parcels were saved as a separate point shapefile.

5. *Counting Parcels by Blocks*. The owner-occupied and total parcel counts of each were tallied for each block. 22 of the 786 blocks contained no parcels and so were removed. The remaining 764 blocks were saved as a separate shapefile.

3.3.3 Merging Owner Records

The 2001, 2003, and 2005 lists of owner-occupied residences were compared by name and address to derive migration patterns at the household level. First, the names were sorted in alphabetical order (last, first). Then the 2001 and 2003 data were arranged side by side in an Excel spreadsheet for line-by-line comparison. When a name appeared in 2001 but not in 2003, the owner was coded as "left" – either he/she left Danville or became a renter. Likewise, if a name appeared in 2003 but not in 2001, the owner was coded as "came" (this could be a renter becoming an owner or a household new to Danville). This line-by-line comparison was conducted manually so that typographical mistakes in the datasets could be readily identified and overcome. Similarly, when the first name entries included more than one name (e.g., Joe & Martha) in one year but one name (e.g., Joe) in another, the entries were recognized as the same "owner" (e.g., the line highlighted in Table 3). The manual identification of matching owner entries in different years was susceptible to human error. However, the lists were processed multiple times so that some mismatches were identified and resolved. The same process was repeated for the 2003-2005 comparison.

Table 3. Example of Owner Matching Across Years[11]

ID0103	NAME	2001 ADDRESS	Code	NAME	2003 ADDRESS
4101			Came	JONES, KIRK	1403 OAK
4102	JONES, LARRY	109 JEFFERSON	Left		
4103	JONES, MADGE	806 SUNSET RIDGE	Stayed	JONES, MADGE	806 SUNSET RIDGE
4105	JONES, MAE	1305 LAPE	Stayed	JONES, MAE	1305 LAPE
4106	JONES, MATT	1202 MARTIN	Stayed	JONES, MATT & ZOEY	1202 MARTIN
4107	JONES, NINA	120 DAVIDSON	Left		
4108	JONES, OLIVER	120 WILSON	Stayed	JONES, OLIVER	120 WILSON
4110			Came	JONES, PAT A	806 GRANT
4111	JONES, PAT J	118 TENNESSEE	Stayed	JONES, PAT J	118 TENNESSEE
4112	JONES, PATTY S	401 ROSELAWN	Stayed	JONES, PATTY S	401 E ROSELAWN
4113	JONES, PETE & DANA	905 HOLIDAY	Stayed	JONES, PETE & DANA	905 HOLIDAY
4118			Came	JONES, RICK & DONNA	735 BRYAN
4119	JONES, ROBERT	220 DENVALE	Stayed	JONES, ROBERT	220 DENVALE
4120	JONES, ROBERT & HELEN	12 BEARD	Stayed	JONES, ROBERT & HELEN	12 S BEARD
4121			Came	JONES, ROBERT L	836 SOUTH
4122	JONES, RON	1205 LOGAN	Moved	JONES, RON	1407 FRANKLIN
4124	JONES, RONNIE & MARTH	716 CLEVELAND	Stayed	JONES, RONNIE & MARTH	716 CLEVELAND
4125	JONES, RUSTY & ANN	1109 CLEARY	Stayed	JONES, RUSTY & ANN	1109 CLEARY
4127	JONES, T J & KERRY	11 NATIONAL	Stayed	JONES, T J & KERRY	11 NATIONAL
4128	JONES, TIMOTHY & DEBR	122 BOWMAN	Stayed	JONES, TIMOTHY & DEBR	122 S BOWMAN
4129	JONES, VERN D	1010 MAX	Stayed	JONES, VERN D	1010 MAX

The process of geocoding was performed by matching the addresses with the parcel GIS dataset provided by the city of Danville and described above.

3.3.4 Census Data Extraction

A significant secondary data source in this project was the U.S. Census 2000. These data align closely in time with the owner-occupied household list assembled in 2001 (which relied upon owner occupancy in the prior calendar year

1. *Collection of Block Data from Census*. Using a query for the 764 selected blocks within the “place” of Danville (see section 3.3.2), Census data were extracted for tenure (housing units owned, rented and vacant) and race by tenure (white householders who owned or rented).
2. *Imputation of Renter-Occupied Parcels*. Because the owner-occupied parcels were already established, the renter-occupied parcels were imputed using the ratio of rental to vacant housing units in the block. This ratio served as a “rental probability” applied to the non-owner-occupied parcels. The quantity of rentals and vacant units was determined in advance, then allocated to parcels in the “Create Model Instances” class, alternating the allocation of rentals and vacant units.

[11] This table is illustrative only. Owner names have been changed to preserve anonymity.

3. *Imputation of Racial Category*. "White" or "Non-white" racial status was imputed on the basis of tenure at the block level.[12] To do so, the probability of a household being white was set to the fraction of white households within the owner-occupied and renter-occupied households. In the "Create Model Instances" class, these probabilities were invoked within the owner and renter algorithms respectively (see Appendix F for the algorithm steps).
4. *Assigning Children and Income*. Census data at the blockgroup level were used to determine the prevalence of children among white and nonwhite households, as well as the distribution of income among white and nonwhite households. Racial status was used to impute children and income attributes because the information was available at that level of detal, and race had already been imputed at the finer-grain block level in step 3. This approach is limited by the fact that blockgroups encompass more blocks than are considered and so the blockgroup distributions may not be representative. By accounting for the cross-distributions of these attributes by race, this approach maximizes the use of available information in hopes that the simulated distributions resemble the actual distributions in Danville.

3.4 Analysis of Geographic Patterns

The availability of owner-occupied household data (see section 3.3.3) enables analysis of patterns appropriate to guide model development and calibration. This section describes two analyses using the geographic distribution of owners: the spatial distribution of households sharing the same last name, and the intra-community migration of households over a two-year period.

3.4.1 Distance Between Households Sharing Last Name

Homeowner last names were included in the records described in section 3.3.3, and were geocoded using the parcel data described in section 3.3.2. This information was used to compute distances between all owner-occupied households in the most recent (2005) record. Distances

[12] The variable code H015I refers to tenure status (rent or own) for non-Hispanic white householders. The racial category "Black or African American" constitute the majority of non-white households in Danville. For simplicity of representation, the racial category was reduced to "white" or "non-white."

between households sharing the same last name were then compared with households having different last names.

Table 4 compares households with shared and unique last names. While the majority of households (72%) share their last name with at least one other households, a very small fraction (0.11%) of the links *between* households occur between households sharing the same name. The vast majority (almost 30 million) of household pairs have different last names.

Table 4. Comparison of Households with Shared and Unique Last Names[13]

	Shared Last Name	Unique Last Name	Total
Number of Households	5628 (72%)	2117	7745
Number of Possible Links	32970 (0.11%)	29,963,415	29,996,385
Mean Distance	9796.2 feet	10236.8 feet	

Table 4 reveals a difference in distance of over 400 feet closer proximity for households with shared last names, relative to households with different last names. These distances were computed using a Python script to convert coordinate pairs into Euclidean distance, and were then imported into R for statistical analysis. The Welch two-sample t-test was performed on the shared versus different last name distances. This t-test revealed a significant difference (t=-14.8071, df=33044.21, p-value < 2.2e-16) of 440.6 feet closer proximity for households sharing the same last name. The 95% confidence interval for this difference in distance was 382.3 to 499.0 feet. The histograms of distance between households illustrated in Figure 15 also reveal that the distribution of distance between households sharing last names is skewed toward shorter distances than households of different last names.

[13] The number of households is listed as the *total* households sharing last names with at least one other household. In contrast, the number of possible links is listed as the connections between households sharing the *same* last name.

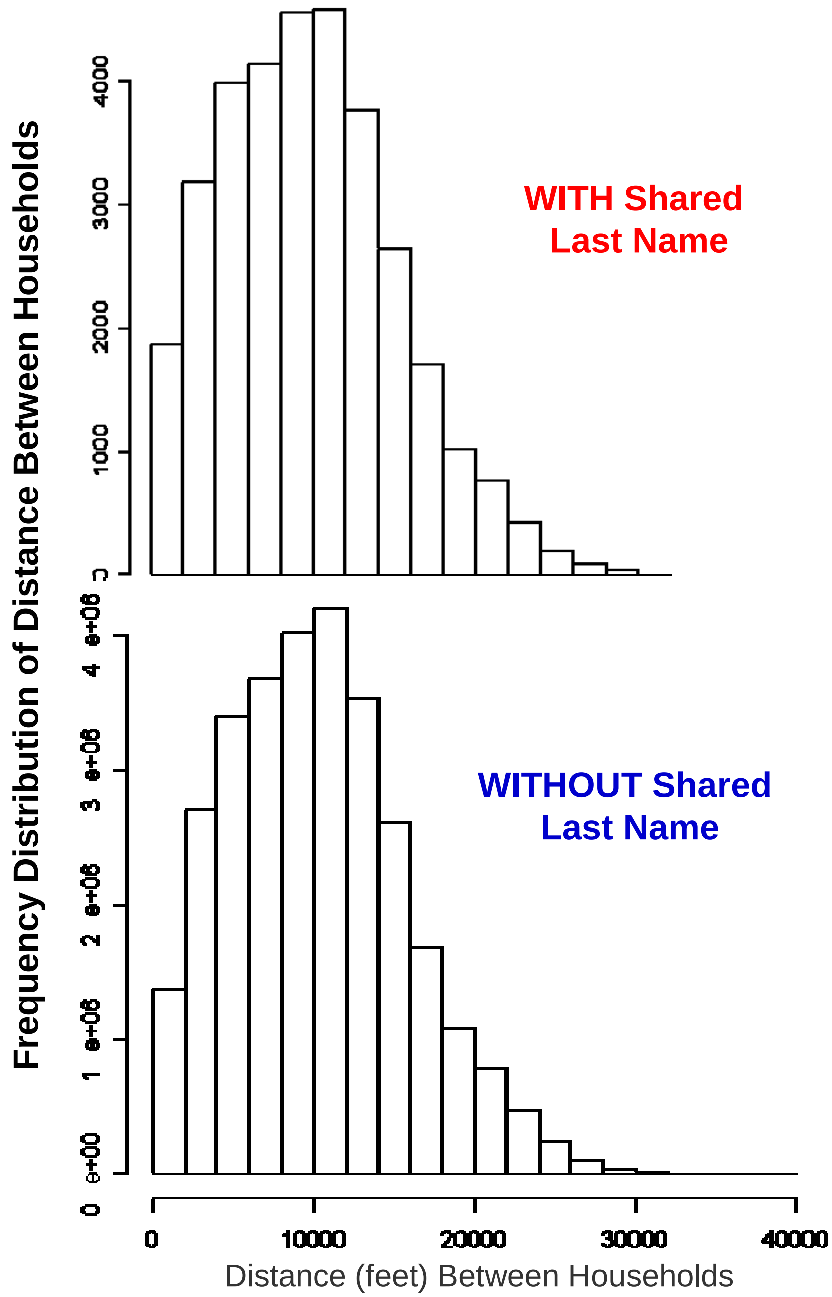

Figure 15. Distribution of Distance Between Households

This difference is more pronounced when the average distance is considered by the uniqueness of the name. The horizontal axis in Figure 16 reveals the commonness of a name as the number of households sharing the same last name (e.g., “2” for two households with the last name of Metcalf). The vertical axis shows how the average distance varies with such commonality of name. The solid line in Figure 16 shows the average distance between households with different last names, which is over 400 feet farther than the distance for those sharing the same last name. However, this average is significantly diluted by the number of households with very common last names (i.e., the "Smith" and "Johnson" effect). For example, the average for households sharing a last name with only one other household is approximately 8500 feet, which is substantially lower than the overall average.

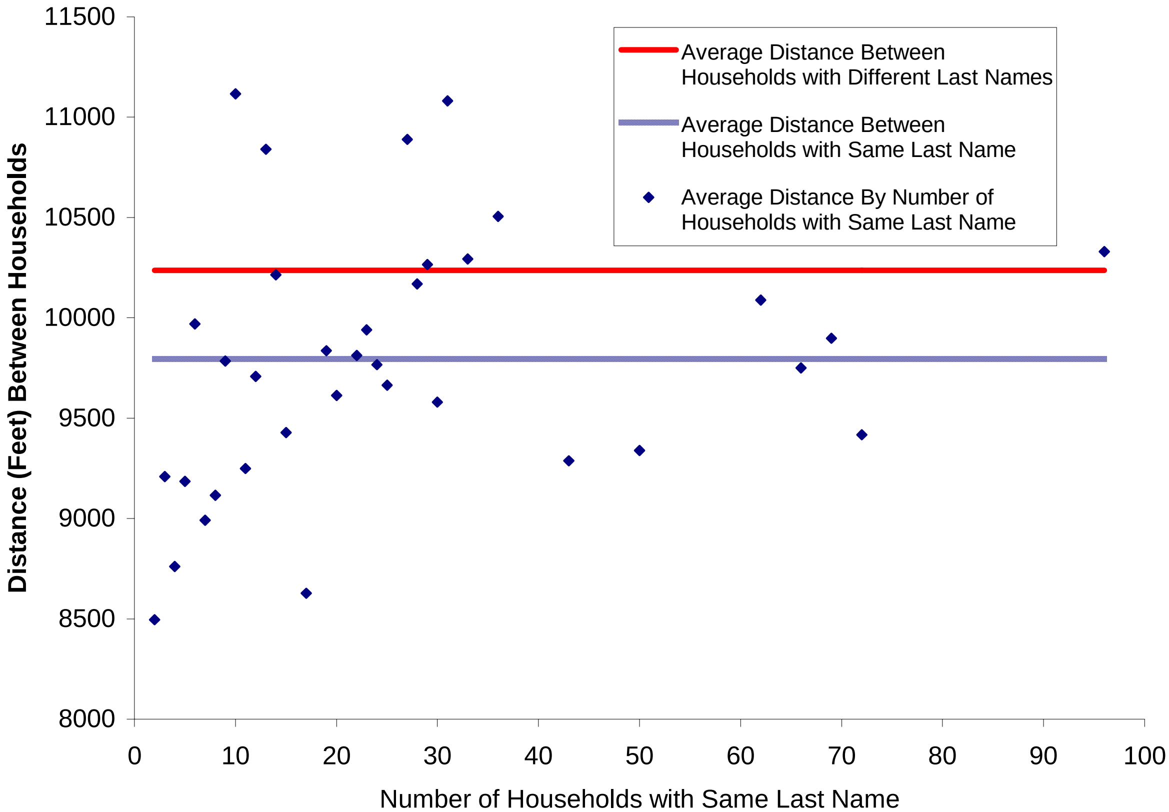

Figure 16. Distance Between Households by Number of Households Sharing Last Names

The dilution of average distance between households by common names stems from the number of possible links between high-frequency names. If a name such as "Johnson" is shared by 90 households, the number of distances to be measuredis 90*89/2 = 4005. In contrast, a name shared by only two households has just one measurable distance.

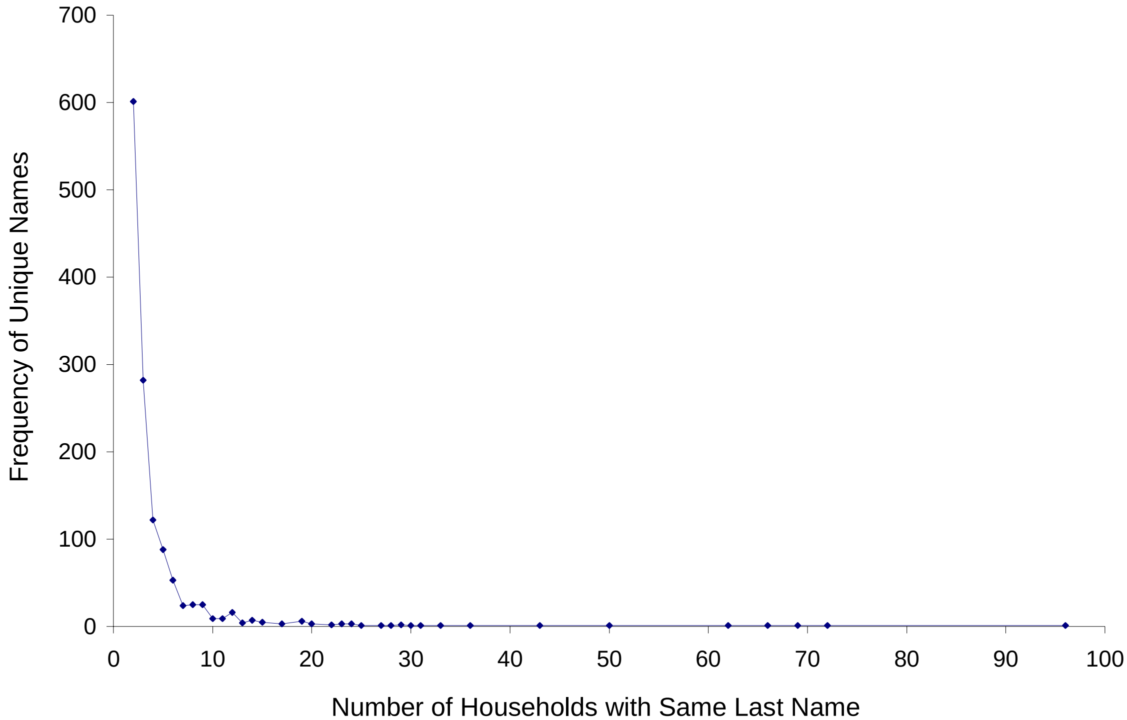

Figure 17. Distribution of Unique Names by Number of Households Sharing Last Names

The frequency distribution of unique names in Figure 17 corroborates the dilution effect that very common last names have on average distance. Approximately 600 names are shared by only two households, and nearly 300 names are shared by three households. This number declines much like a power law, which has come to be a signature measure for complex systems, such as the degree distribution of scale-free networks (Barabasi and Albert 1999).

The analysis of distance between households sharing last names provides a means of testing an implicit social network effect. Two obvious limitations of this analysis are that: 1) not all households sharing last names are related, and 2) not all sets of relatives in Danville are reflected due to their different last names as well as the exclusion of renters from the source data. These limitations render the analysis difficult to embed in a simulation model, although a consideration of both last name *and* uniqueness could strengthen the social network effect on migration choices that result in patterns of disparity. As described in subsequent chapters, this option is reserved for future extensions of the model.

3.4.2 Creating the Owner Migration Pattern

The most important part of the modeling process illustrated in Figure 5 earlier is the creation of patterns both to guide model development and to provide points of comparison for model analysis. This section describes the creation of owner migration patterns from the time series homeowner data described in section 3.3.3 above.

Parcel centroids were used to create point feature sets, isolating only the owners who moved in each two-year time period. The from-to pairs of points were then used to create line features as vectors showing migration paths of individual households over two years.

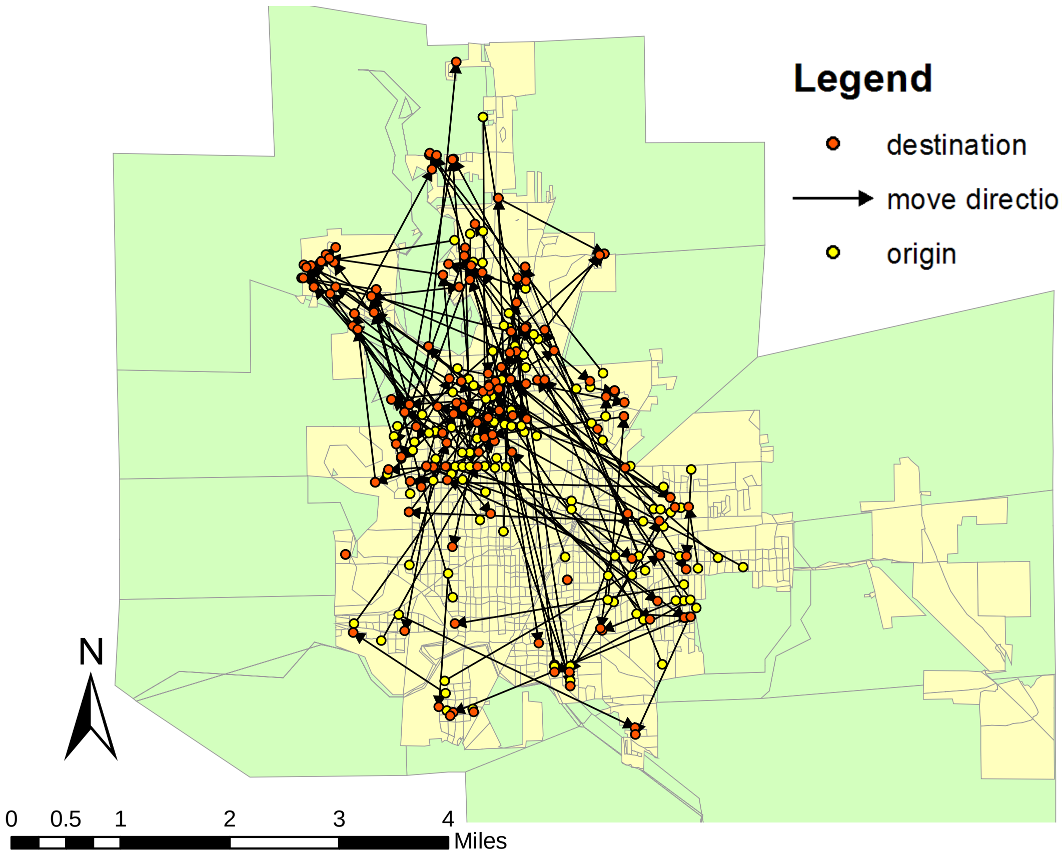

Figure 18. Migration Within Danville from 2001 to 2003

The set of overlapping arrows in Figure 18 represents vectors for homeowner moves between 2001 and 2003, derived from the name and address information provided by the city of Danville. The light dots represent the 2001 origin, and the darker dots represent the 2003 destination. While the entire set of arrows in Figure 18 makes directionality difficult to distinguish, patterns may be distinguished by selecting subsets of parcels. Selecting parcels in the center reveals a strong tendency to point to the northwest direction, which is the affluent section of Danville. The desirability of the northwest neighborhood is confirmed by selecting the parcels in the northwest section by the lake – most of the moves are to the area, and not away from it. Selecting parcels in the lower east side (i.e., ward 3 in Figure 12 where the interviews were conducted) reveals a strong northwest outmigration tendency, though not reaching as far as the lake. An internal migration pattern for Danville would be this northwest vector. For the purpose of calibration, these changes were tabulated at the blockgroup level (see Appendix E).

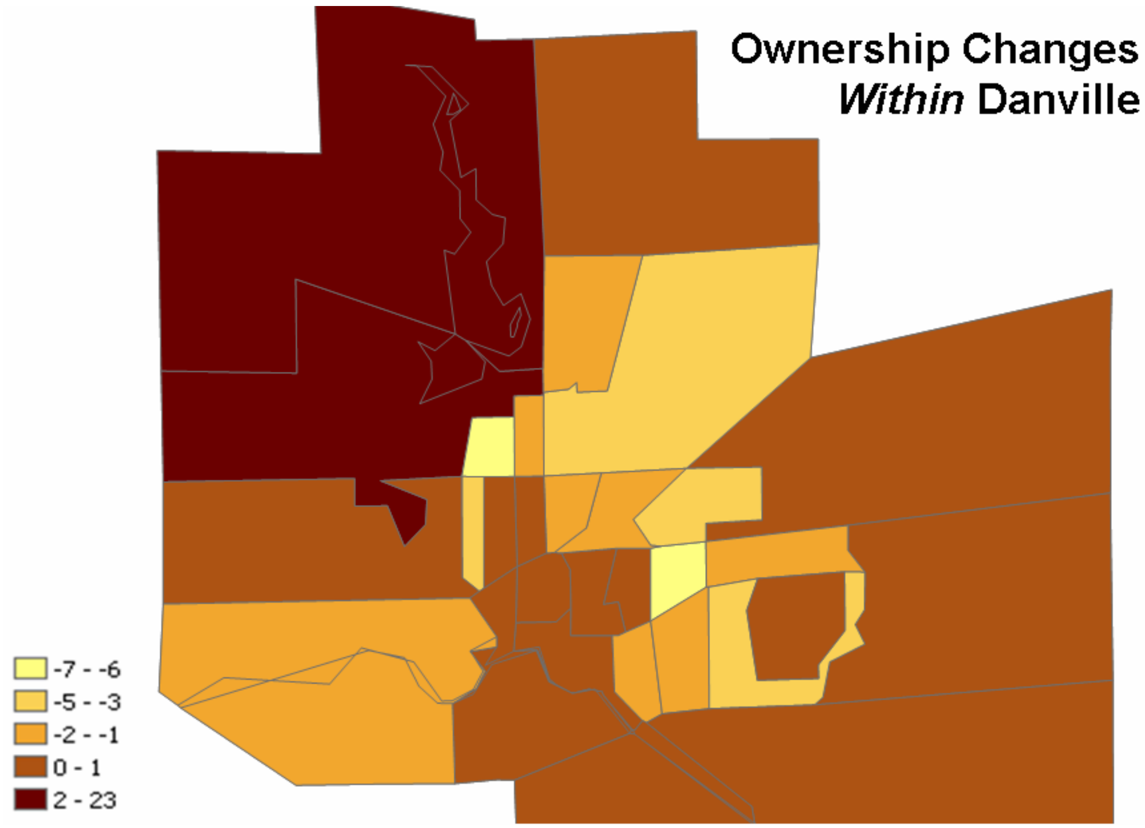

Figure 19. Ownership Changes From 2001 to 2003 Within Danville

These ownership changes within Danville reveal internal migration preferences. Figure 19 reveals the aggregate effect of the individual arrows shown in Figure 18. The dark areas, such as the lake area, have the most incoming owners, while the lighter areas represent areas of outflow.

However, for the same time interval of 2001 to 2003, if *all* the ownership changes – new owners as well as owners who left Danville – are included, the pattern nearly inverts (Figure 20).

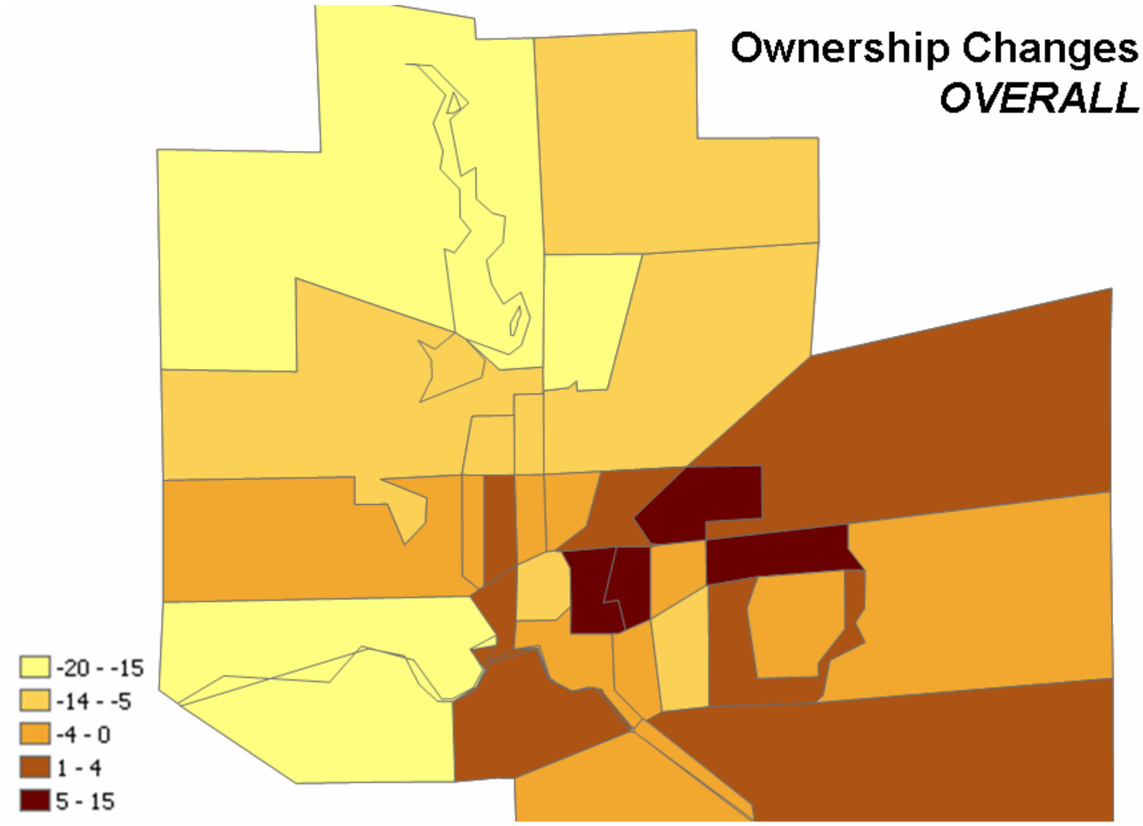

Figure 20. Overall Ownership Changes From 2001 to 2003

For this time period (2001-2003), Danville experienced a net outflow of 55 owner households, many of whom left the attractive northwest lake area. Even more confusing was the inflow of owners to the southeast side of town, that seemed less attractive. The key to the explanation for the behavior exhibited in Figure 19 and Figure 20 was that an increase overall included both newcomers to Danville *as well as* renters who become owners. It is not uncommon for landlords (once they've made enough money from renting) to sell houses to tenants on contract. In my neighborhood of focus in the southeast side of town, there is a housing investor who buys a house to refurbish it, find a good renter, and help them transition to ownership.

Apparently this practice is common in low-moderate income areas (often areas that are implicitly red-lined[14] because banks won't lend readily there) and less common in higher-income areas where "normal" housing market dynamics are at play. So the overall increase in ownership in lower-income areas may reflect a significant number of renter to owner transitions.

[14] Red-lining is a form of mortgage discrimination in which certain areas, once literally marked on maps with a red line, are considered off-limits to receive bank investments. This practice is now illegal in its explicit form.

The net exodus by the Lake area (despite its appeal to movers within Danville) may be accounted for by the link between affluence and mobility. Affluent professionals are able to afford much more than Danville has to offer in terms of housing investments. Danville is a depressed housing market, as revealed in a recent survey with its median home price of $65,900.[15] The only parts of Danville where housing investments can be recouped are the wealthiest (by the lake, ward 7 in Figure 12). Professionals move there but frequently have to move elsewhere for work. Moreover, there are several satellite communities around Danville that are destinations for some movers. But in the lower income areas, the housing values are considered to be already as depressed as they're going to get. This allows people to not worry about losing investments, as long as it's a sound structure. One of the residents interviewed for this project had this perception as a new mover. This resident felt confident with the housing value not declining further, but decided not to invest in a garage addition because the value of such an addition would probably not be recouped in a later resale.

For the 2003 to 2005 ownership changes, the broad pattern remained the same – within Danville, the northwest (lake) area was attractive; but overall changes demonstrated a net outflow from attractive areas and an inflow of new owners to lower income areas. In summary, the people who live in the most affluent areas were most able (and likely) to move elsewhere. Renter to owner transitions were more common in lower income areas, as well as new movers who want a good investment but can't afford the lake area.

This explanation of overall migration patterns suggests the importance of economic forces, modeled as an affordability threshold in the next chapter. Perhaps with higher income one can “escape” an undesirable neighborhood, but with a lower income one locates closer to the valued social connections. Indeed, several residents of the lower income neighborhood of focus for this project were family members co-locating to provide child support and other valued social services that are not readily measured by household income.

While the discrepancy between intra-Danville migration patterns and patterns that include the overall inflow and outflow is highly intriguing, the ambiguity about causes of inflow and outflow render the overall pattern less useful for model calibration. The intra-community ownership migration pattern is therefore used for the calibration results described in Chapter 5.

[15] “Housing Report for Danville.” 11 November 2006. http://www.ecanned.com/IL/Danville_MSA.shtml

CHAPTER 4. MODEL DEVELOPMENT

This chapter describes how an agent-based model was constructed to simulate spatial and dynamic household choices over time. Before building a complete model for the City of Danville, a prototype model (Metcalf and Paich 2005) was constructed based upon an abstract two-neighborhood conceptualization. Then, parcel and Census data were integrated into a robust modeling framework to simulate dynamics of approximately every household in Danville.

AnyLogic[16] simulation software was used for both model development and calibration (as described in the next chapter). Based upon the Java object-oriented platform, AnyLogic uses graphical icons to represent modular objects and algorithms. While the AnyLogic environment was helpful for model conceptualization and simulation, most of the data integration steps took place directly in the Java programming language using the Eclipse integrated development environment (IDE).

4.1 Model Conceptualization

One of the first tasks in developing a model is to visualize it (Grimm and Railsback 2005).The template in Figure 21 demonstrates how a neighborhood network may be embedded in two-dimensional space. In this case, houses are distributed uniformly across a grid between two neighborhoods. Social links are strong in broad domains ("neighborhoods"), though not restricted to direct neighbors (as in Schelling 1971, 1978). This diagram is a juxtaposition of the relational space of the social network as visualized in two-dimensional space. The social links of one house from each neighborhood are shown. While links are predominantly contained within each house's neighborhood, they expand beyond direct neighbors and may ultimately connect across neighborhoods through boundary households.

In this abstract two-neighborhood space, households may have spatial preferences for friendship connections in their immediate neighborhood or in proximate neighborhoods. Or maybe social selection preferences dominate. Income threshold may be considered as a proxy for socioeconomic status for starters, though other factors could also be included. And then, the impact of moving is considered: Does a household change social connections after a move? Is

[16] http://www.xjtek.com

there a difference when a household moves between neighborhoods versus out of the community altogether?

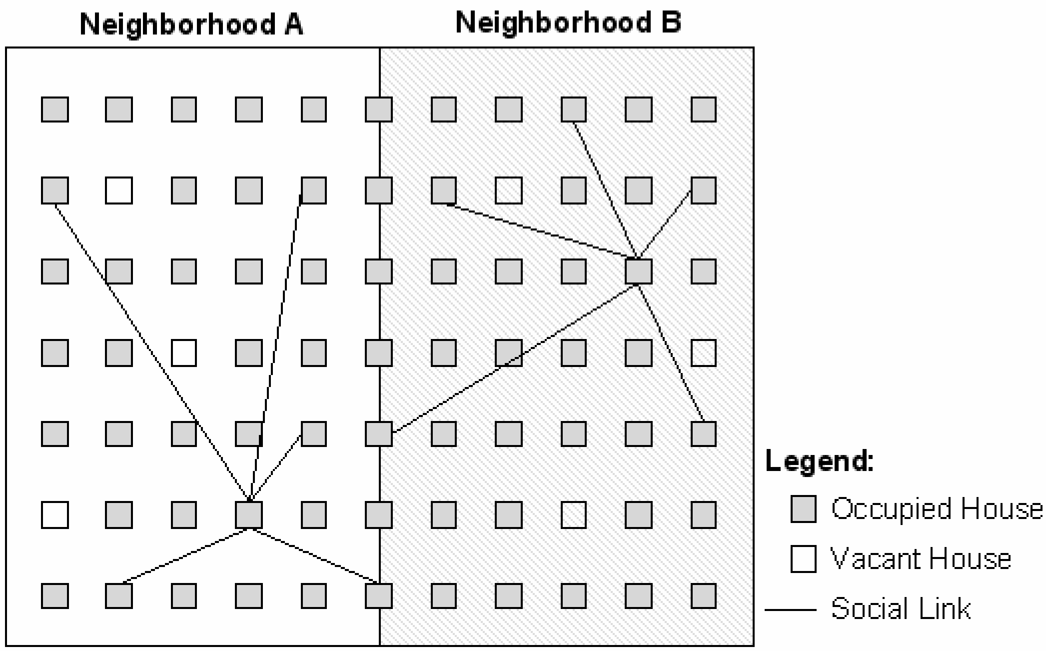

Figure 21. Example Template for a Spatial Social Network

This simple conceptualization is the beginning of an agent-based model. These households are *agents* that connect with other households both inside and outside the neighborhood. They may also move between neighborhoods, and into or out of the community altogether. An abstract example model structure may be built from these heuristics.

Figure 22 is a screenshot of this prototype model as developed in the AnyLogic modeling environment. The left panel shows the model objects: house, neighborhood, person (in this case, one person created a household), and *main*. An agent is an object, but an object is not necessarily an agent. Objects may undergo state changes without making decisions per se. As described earlier, the AnyLogic environment is built on the Java object-oriented programming language, making it amenable for agent-based modeling. Featured in the center panel, the *main* object is the "view of God," as it contains all of the model – multiple instances of the state objects neighborhood and house, as well as the "people" agents. It also contains aggregate indicators such as cohesion and income gap between neighborhoods. At the right is a view of the animation of this model just after it starts up. This is not a particularly informative or aesthetic visualization, but it does reproduce the two-neighborhood structure just articulated.

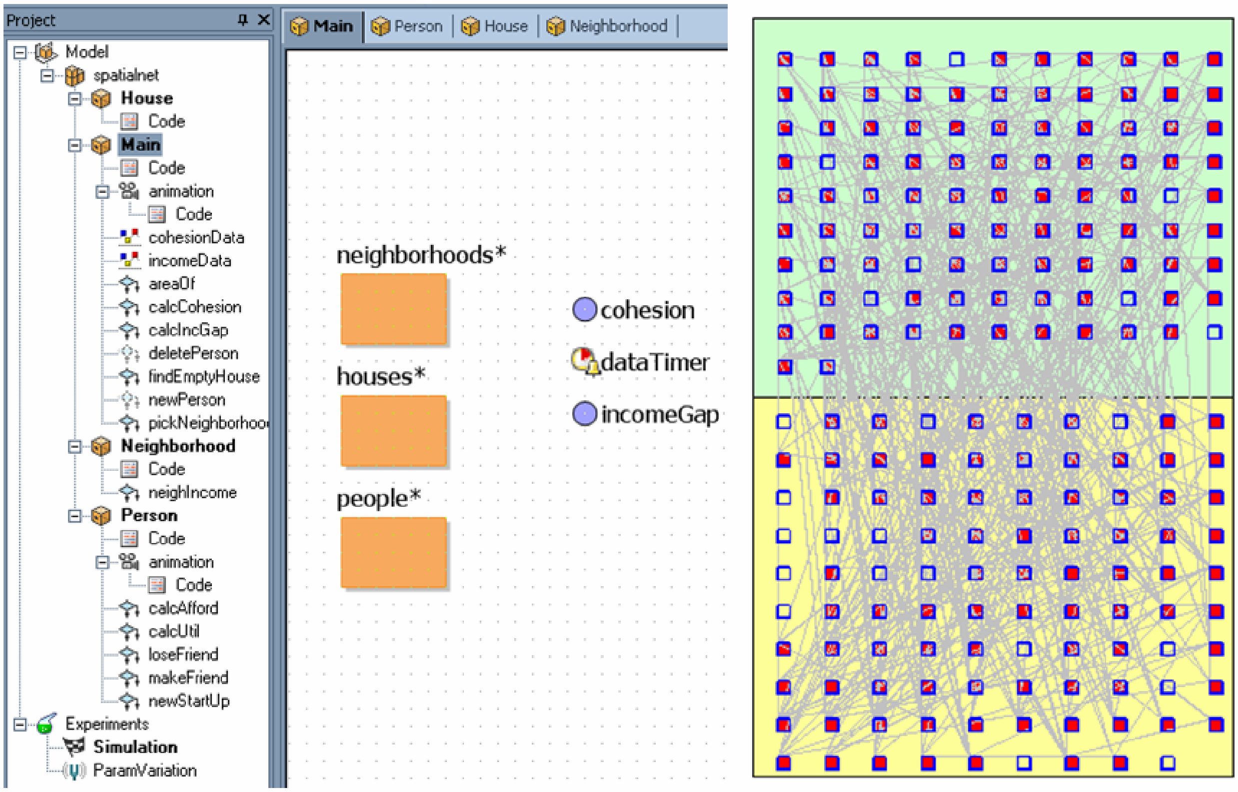

Figure 22. Screenshot in the AnyLogic Simulation Environment

The dynamics of this prototype model depend upon decision rules made at the individual level. Two choices were made: who to like and where to live. Both choices employed a logit expression of the utility of weighted effects. Neighborhood utility derived from income-based attractiveness and the individual's fraction of friends in the target neighborhood. While neighborhood attractiveness was a universal measure (proportional to the neighborhood affluence relative to overall community affluence), the neighborhood-based friendship fraction was unique to each individual. In turn, friendships reflect spatial and socioeconomic preferences through the utility of a social network connection.

A stochastic, exponential distribution of decision evaluation frequencies enabled asynchronous evaluation by individual agents. Whether or not a move was made depended upon the neighborhood utility, and the availability of houses in the target neighborhood. If a move was made, the relocating agent updated their friendship network by making one friend and losing a friend according to a binary logit of utility.

Experiments with this prototype model (Metcalf and Paich 2005) revealed that the income-based attractiveness is stronger in the utility equation when neighborhood preferences are balanced. This occurs because the neighborhood friendship fraction averaged 50% (equal base probability of choosing a friend in either neighborhood) in the balanced case, versus 70% (stronger preference for local neighborhood) in the unbalanced case. Those agents who could afford (based upon their income level) to move to the more attractive neighborhood did so, producing an agglomeration of higher income individuals in the attractive neighborhood. At the same time, balanced preferences resulted in a higher cohesion (fraction of social ties that cross neighborhood boundaries) level. This co-occurrence of balanced (a-spatial) friendship preferences and community-wide network connectivity in the prototype model demonstrated that spatial disparity could result without social network fragmentation (the converse of cohesion) under this formulation.

The prototype model was used to explore how the concepts of neighborhood and network choice could be combined in an algorithmic agent-based model. While additional insights could be gleaned from testing this abstract model, it did little to verify the adequacy of assumptions relative to a real-world case study. The model as developed for the case of Danville, Illinois required substantial parameterization based upon the data sources outlined in Chapter 3. This model development is described in the sections that follow. Structurally, the major differences from the prototype model are the extent to which location is accounted for more explicitly (at the Census block level), and the inclusion of parameter weights corresponding to Census-derived attributes that then may be adjusted as alternative model structures. The following sections describe how the model structure develops from a set of data objects for the case of Danville.

4.2 CaseTown Model Structure

Moving from an abstract prototype to an applied model required significant consideration of what the applied model structure contains. At the heart of the model, household agents make decisions about where to live (neighborhood) and who to associate with (social network) based upon factors of geographic proximity, income, race, and presence of children. These dynamics take place in a landscape of parcels that may be owned, rented or vacant, and are spatially situated within Census blocks and blockgroups for the city of Danville.

The initialization of the household agent objects and the spatial state objects was a significant step in structuring the model. This section outlines the structure of the applied

"CaseTown" model as informed by Danville data. Supporting documentation is provided in Appendices D (raw data files), F (java classes), and G (AnyLogic implementation).

4.2.1 Creating Model Objects from Data

As noted in Chapter 2, an agent is an object but an object is not necessarily an agent. For the full model, households are the only true agent objects, embodied as mobile agents that make decisions about where to live and who to include in their social network. The household sets the scale of decision making; it is not further subdivided into family members. Households may own or rent parcels, which are plots of land corresponding to location data provided by the city of Danville. The parcel object is a state object, *not* a decision-making agent. Likewise, blocks and blockgroups are state objects representing census areas of aggregation. Census blocks inform the rent/own/vacant status, and the blockgroups inform the income distribution of the households. The spatial scope of the block objects is revealed in Figure 23.

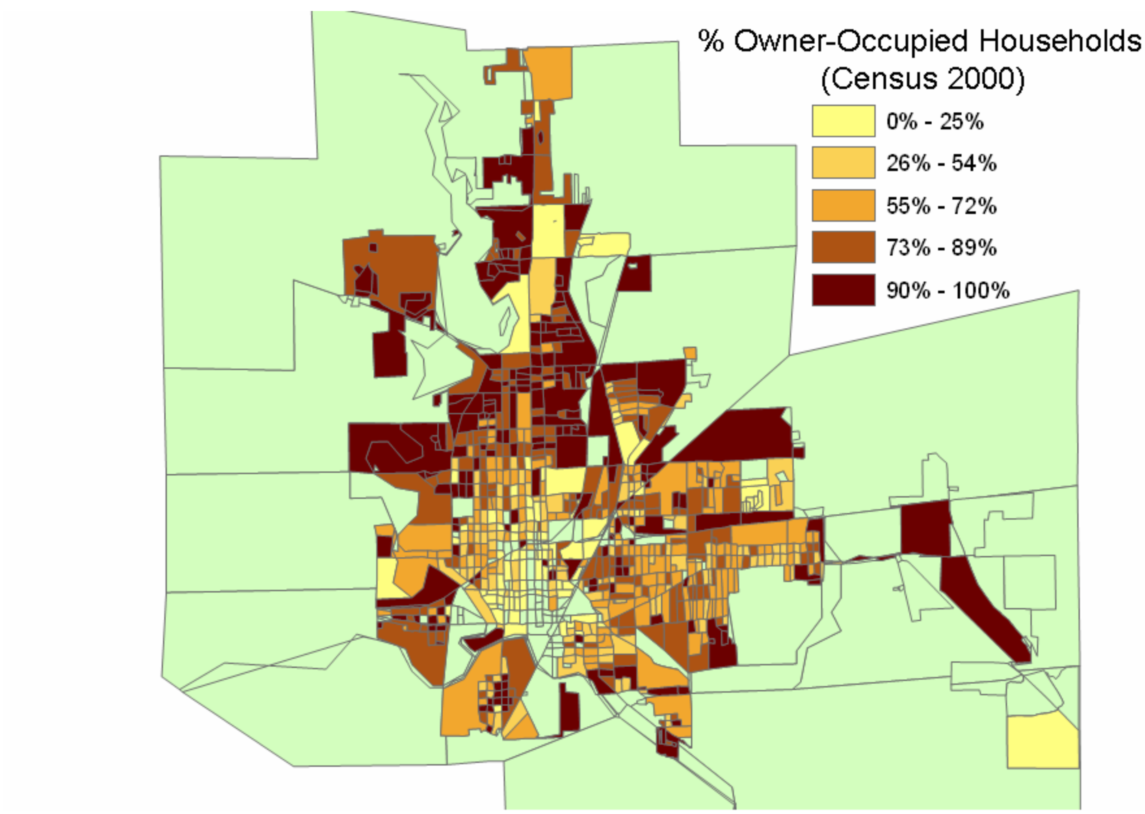

Figure 23. Map of Owner Occupancy in Danville Census Blocks

The map in Figure 23 illustrates the prevalence of owner-occupied households in census blocks. The light backdrop illustrates the larger blockgroup boundaries. The darker shades indicate that owner-occupied households dominate renter-occupied households. The percentage shown does not include vacant households, so that it is strictly a ratio of owner-occupied to all occupied households. Danville city officials consider the renter-owner ratio to be a significant determinant of neighborhood stability, and resident interviews corroborate this hypothesis. Because of this importance, rent and own states are incorporated into the model at the parcel level.

After defining what the model objects represent, the non-trivial task remained to create the objects from Danville data. The schematic in Figure 24 illustrates the process of creating model objects. This process began at left with unstructured, or raw, data files. Unstructured does not mean unorganized – it simply means that the data lack semantic meaning. The geographic data processing steps described in Chapter 3 enabled the creation of four simple text files, where the rows correspond to the number of objects. 28 blockgroups contained the income and move frequency data, 764 blocks contained the rent-own-vacant information, 13,166 parcels contained location coordinates, and 7,576 initial owners contained parcel identification. Table 9 (parcel), Table 10 (owner), Table 11 (block), and Table 12 (blockgroup) in Appendix D provides excerpts from these four raw data files.

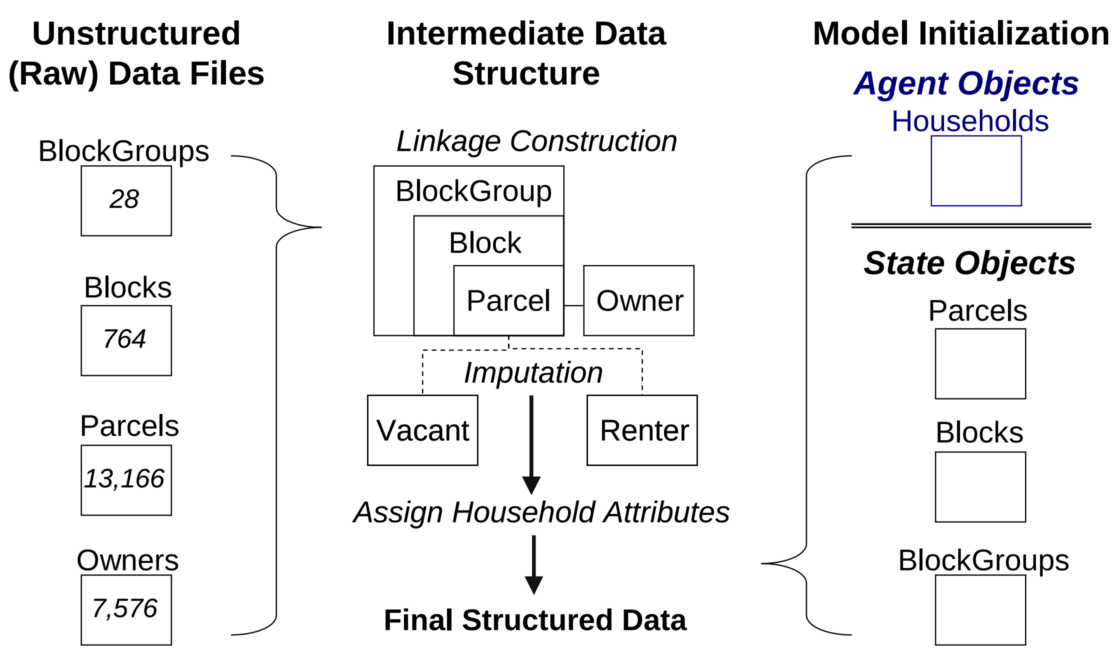

Figure 24. Process for Creating Model Objects

As outlined in Figure 24, the unstructured data were imported into the Eclipse development platform for processing. Apendix F documents the sequence of tasks for these data to become structured. The first task was to properly link the data. Spatially, these links dictated that blockgroups contain blocks, which in turn contain parcels. Parcels were linked to their initial owners (from the 2001 owner-occupied household list) as described in Chapter 3.

After the structured data were properly linked, the remaining objects and attributes were identified. Section 3.3.4 describes how rent and vacant parcels are imputed from block characteristics, and how household racial status was imputed from the own or rent status. Then household income and presence of children were assigned based upon blockgroup characteristics and household racial status.

The java classes in Appendix F provide explicit documentation for the creation of initial model objects from raw data. The sequence of processing classes was as follows:

1. Create Data Structure (args: 28blockgroups.csv, 764blocks.csv, 13166parcels.csv, 7576owners.csv, dataStructure.ser). This class created data containers to provide structure to the raw data.
2. Create Model Instances (args: dataStructure.ser, modelObjectsFinal.ser). This class imported the structure created in the previous step, and created the template for the model objects, including the assignment of rental and racial (boolean, true for “white”) status.
3. Block Distance (arg: modelObjectsFinal.ser). This class imported the newly defined set of model objects and updated it to include a block distance matrix for the N=764 blocks. This created a matrix with 292,230 distance measures (see **Error! Reference source not found.**). The maximum distance was 34,737 feet and the average distance was 9397 feet between blocks.
4. Assign Children & Income (arg: modelObjectsFinal.ser). This class imported the latest instance of model objects, and assigned children (boolean, true if present) and income (from a distribution across 16 income categories, bounded by 5k and 250k) attributes based upon household racial status at the blockgroup level.

4.2.2 AnyLogic Model Structure

At this point the final structured data had been written to a file, and were ready to import into the AnyLogic simulation environment as model objects. The distinction between agent objects and state objects is important to this structure. Households are true agents, making decisions that change their state and the state of other objects. Parcels, blocks, and blockgroups are objects that change state but do not make decisions.

Figure 25 illustrates a screenshot of the model structure in AnyLogic. At the upper left are the state objects Parcel, Block, Blockgroup, Household, and ModelObjects. The Household state object is a container for state information used to create the dynamically active Household agent object. The ModelObjects class is a container for the sets of objects (encoded as java ArrayLists) blockList, householdList, etc. The CaseTown package in Appendix F provides full documentation for these classes.

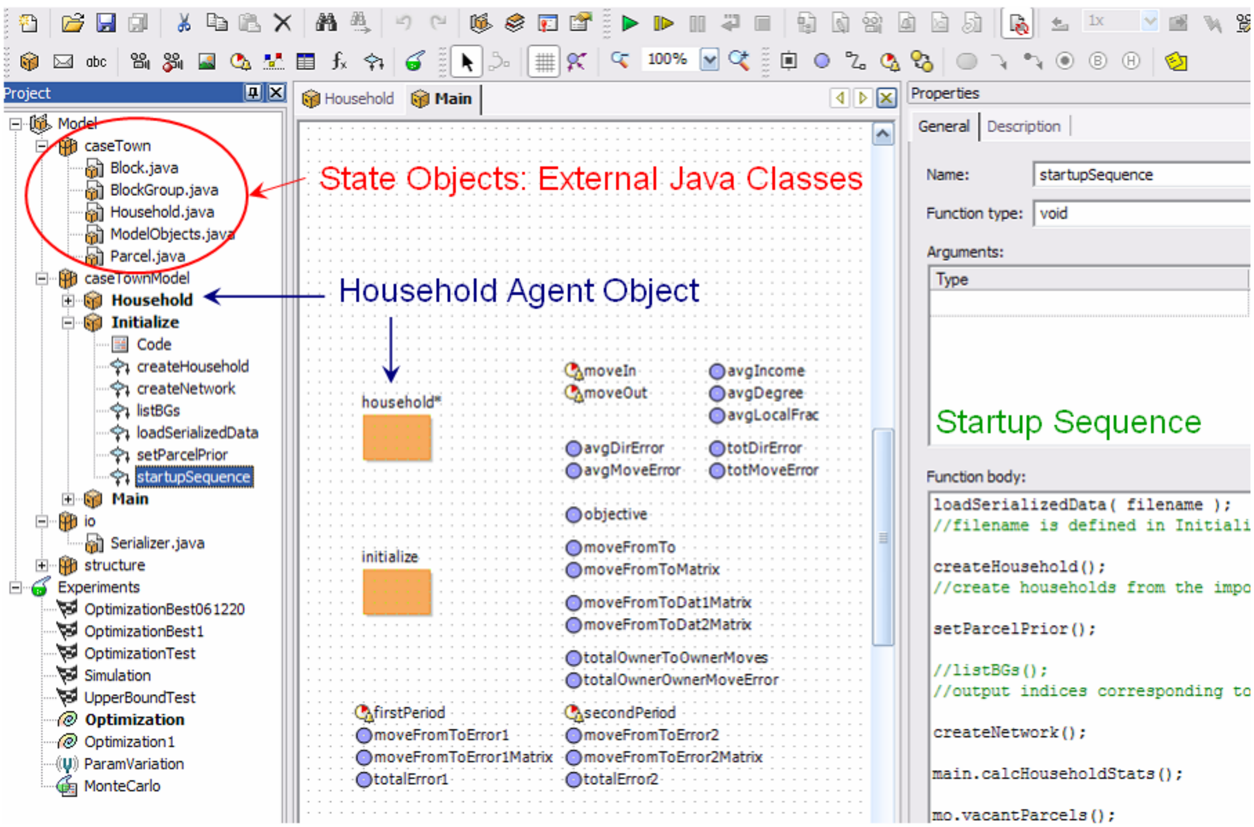

Figure 25. Model Structure Screenshot in AnyLogic

The center panel of Figure 25 shows the structure of the main class in AnyLogic, which (as mentioned above) is a sort of "view of God" encompassing all of the other class objects. The AnyLogic references "Active Objects" the dynamic objects whose instances are created wholly in AnyLogic. This includes the Household agent object as well as an Initialize class that contains the startup sequence illustrated at right. This startup sequence is executed as the first line in the Main class startup code (see Appendix G).

The first step in the startup sequence is to load the serialized data that were created using the sequence of java processing classes described above and documented in Appendix F. These data are then represented as state objects in the caseTown package, using the ModelObjects class container to "hold" the collections of objects. Once the data have been opened in AnyLogic, the Household agent objects are created from the imported Household state objects by iterating through each Household in the ArrayList (see Appendix G for the Initialize class algorithms).

The next significant step in the startup sequence is to create an initial network of household agents. This is accomplished by cycling through the set of agents and requiring a minimum number of connections to be made by each household. The mutual nature of connections ensures that most households are chosen by others before they make their own choices. Here, a minimum requirement of 3 connections results in an average degree of 6 connections. The specific degree value varies with each initialization due to the random component of household selection.

While households pick other households at random for consideration, the choice of whether to make a friendship is not random (although probability is involved). The choice structure for social networks is described below. The factors in households' choice of social ties include distance between blocks, income similarity, racial similarity, and presence of children, as well as scaling constants. A parameter specifying a fraction of friendships to occur within the household's blockgroup is also included. These parameters are adjusted during model calibration as described in Chapter 5.

4.3 Model Dynamics

With the model structure in place, the dynamics were developed. In this section, an overview of dynamics is first provided using a stock and flow representation (as introduced in Chapter 2) emphasizing the importance of feedback and accumulation, which creates delays in the system. After this overview, the specific formulations for the dynamic choices of networks and neighborhoods are presented. All details of the AnyLogic model structure and dynamic algorithms are documented in Appendix G.

4.3.1 Stock and Flow Representation of System Dynamics

The causal map in Figure 26 illustrates the significant forces underlying migration patterns in a place like Danville. To the medically untrained eye, the diagram in Figure 26 may resemble the structure of a heart. Perhaps this visual illusion engenders the empathy appropriate for this study of spatial disparity.

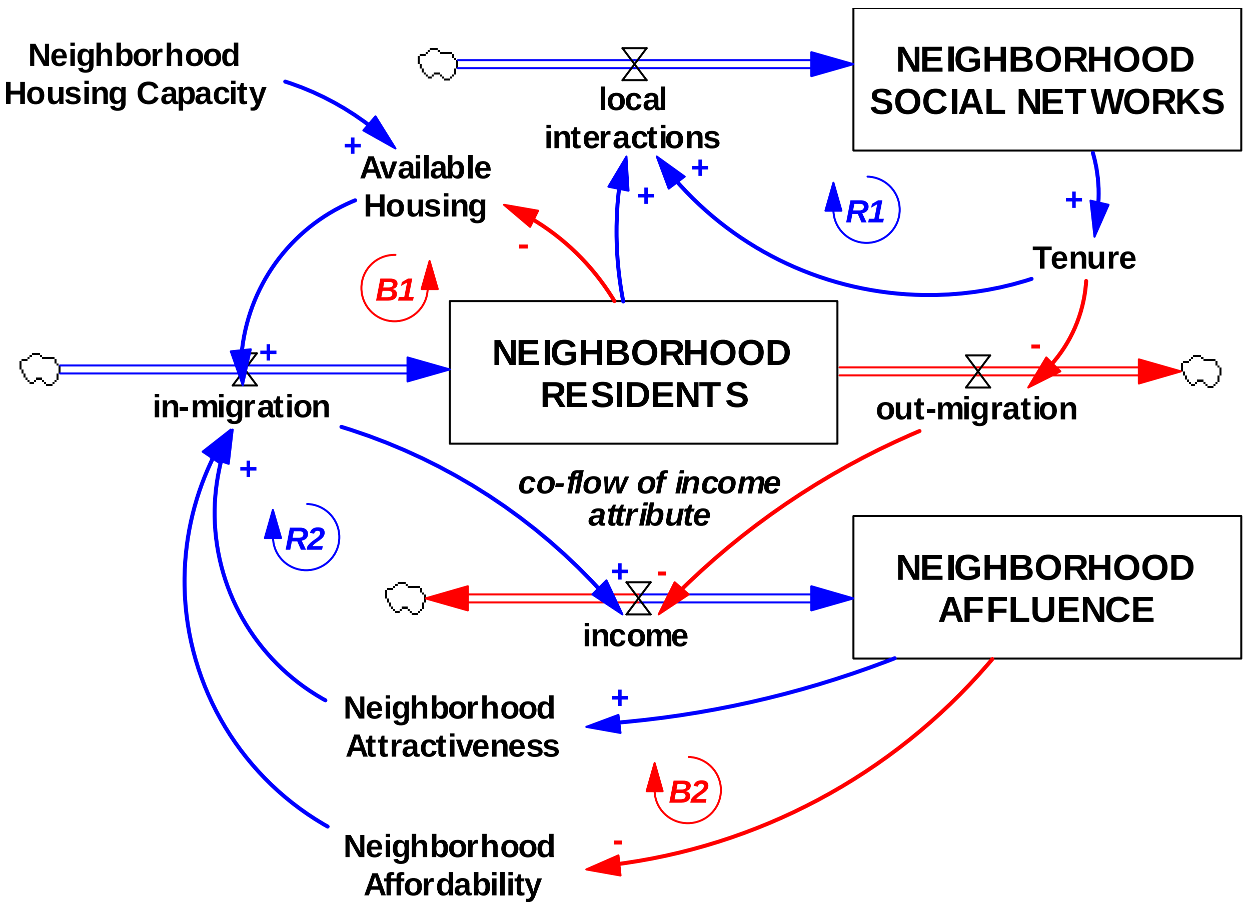

Figure 26. Causal Map of Neighborhood Dynamics

The representation of these system dynamics utilizes the stock-flow iconography that is described more fully in Appendix H. The directionality and polarity of causal relationships is indicated by arrows that create reinforcing (R) and balancing (B) feedback loops. Three major accumulations (stocks) are represented in the boxes. First, the stock of residents changes with rates (flows) of in and out-migration to and from the neighborhood. As the number of residents increases relative to a fixed neighborhood housing capacity (in the simulated case of no new development), a balancing feedback loop (B1) is completed as the number of vacancies decreases. Another effect of increasing the number of residents, in conjunction with the duration of one's tenure in the neighborhood, is to increase opportunities for local interactions. Such interactions may occur within deliberately formed neighborhood associations, or simply through casual contact. These interactions, integrated over time, form neighborhood social networks. As these networks are strengthened, increased duration of tenure in the neighborhood is likely to result, reinforcing (R1) the network effect further.

A "co-flow" representation in Figure 26 enables the flow of incomes associated with migration patterns to contribute to neighborhood affluence, which feeds back to in-migration through affordability (B2) and attractiveness (R2) effects on neighborhood choice. The form of these effects are described in the next section. Note also that this feedback implicitly occurs across scale as neighborhood affluence contributing to the household's choice of where to live.

4.3.2 Decision Rules

How does a household make decisions? As revealed through the ethnographic fieldwork, the social network influence may be strong, but it is also subtle. As modeled, the social influence consists of a fraction of friends within a specified blockgroup boundary encompassing the place (parcel) of consideration. In the real world, the causal mechanism varies from individual to individual. For example: divorce, a marriage, a friend selling property, and a family member passing away are a few of several social mechanisms for residential relocation. Closer to the model assumptions is the influence of discretely bounded neighborhood associations. These association effects are partially inscribed as the fraction of all social ties that fall within a neighborhood.

As noted in Chapter 2, the implementation of agent-based models generally involves algorithms. It has been a significant effort simply to articulate the algorithms. In this way, the modeling practice aids articulation of hypotheses of human behavior. At the individual level, how can we represent (articulate) social influence on neighborhood choice? Here, rules are embedded for social network choice. The resulting fraction of friends in the neighborhood is then measured as an influence on the household's decision of where to live.

The question asked in the pattern-oriented modeling approach is: for the assumptions made, how close do we get to reproducing observed patterns? As noted by Agarwal et al. (2002), the decision-making dimension of model complexity is frequently shallow when compared to the spatial and temporal dimensions. The focus on human decision-making in this model is intended to add depth to this dimension of complexity. What would happen *if* we all thought a particular way? It doesn't have to be the same way; we can have heterogeneity of individual preferences (represented here as a happiness threshold) as well as different rules altogether. In this model, different rules are represented by different weights on parameters in a set of influences on two household choices.

Household agent objects make two decisions: social network choice and neighborhood choice. The social network choice in Equation 1 employs a binary logit expression, such that the probability of a household *i* connecting with household *j* is based upon the exponential of the utility of that connection, divided by 1 plus the same exponential term (Ben-Akiva and Lerman 1985). In turn, the utility of the connection is expressed as a set of alpha (α) parameters multiplied by effects.

Equation 1. Social Network Choice

$$P_{ij} = \frac{\exp\left(U_{ij} / b \right)}{1 + \exp\left(U_{ij} / b \right)}$$

$$U_{ij} = C + \alpha_D \cdot D_{ij} + \alpha_I \cdot \frac{\left| I_i - I_j \right|}{I_{avg}} + \ldots$$

where

P_{ij} = Probability of household *i* to connect to *j*

U_{ij} = Utility of connecting household *j* to household *i*

D_{ij} = Distance between household *i* and household *j*

I_i, I_j = Income of household *i* and household *j* respectively

I_{avg} = Average income of all households in the Danville model

C = Constant average utility

b = Scaling constant

α_D, α_I = Parametric weights for the distance and income effect respectively

The two effects shown in Equation 1 are the distance effect and the income effect. In both cases, greater difference decreases the likelihood of connection. Thus, the weights for these effects, represented by the alpha terms, would be negative when making a connection and positive when breaking a connection. The *C* term represents the constant average utility. This term, as well as the alpha parameters, will be adjusted as part of the model calibration. Additional terms may also be included to create alternative model structures. These terms include binary (dummy variable) representation of race (white or nonwhite) and children (present or not). If the prospective household has the same race, the value of the term is 1. Similarly, if both households have children, the binary term is encoded as unity. If one or both households does not have children, the effect is zero.

Household agents evaluate their social network at a stochastic frequency designed to create asynchronous decisions – which enables realistic heterogeneous behavior, and is also more efficient computationally. At this point in time, a household picks another household at random and tests whether the probability from the binary logit passes a "satisfaction threshold", which is a proxy for human idiosyncrasy.

For neighborhood choice, the algorithm is similar to that of social network choice. The utility of a parcel *p* to household *h* invokes a constant *C* term and a set of weighted effects, where the alpha parameters are the adjustable weights.

Equation 2. Neighborhood Choice

$$U_{p,h} = C + \alpha_A \cdot A_{p,h} + \alpha_I \left(\frac{I_{b(p)}}{I_{avg}} \right) + \alpha_{SN} \left(\frac{SN_{h,bg(p)}}{SN_h} \right) + \ldots$$

$$A_{p,h} = \max \left(\frac{I_h - I_{b(p)}}{I_{b(p)}}, 0 \right)$$

where

$U_{p,h}$ = Utility of parcel *p* to household *h*

$A_{p,h}$ = Affordability of parcel *p* to household *h*

SN_h = Social network of household *h*

$b(p), bg(p)$ = Census block, blockgroup containing parcel

The first effect in Equation 2 is an affordability constraint, such that the effect is negative if the household's income is less than the average block income of the parcel. If the household's income is sufficient, affordability is not a constraint. For assessing the utility of the current location, affordability is excluded from the evaluation. The next effect measures attractiveness based upon the average income of the parcel's block, normalized to the average income of the entire community. Then a social network effect is included. This is simply the fraction of a household's social network (or number of connections) that reside in the destination blockgroup, divided by the household's complete social network. Additional effects may be added to create alternative structures.

Household agents evaluate potential parcels at a frequency informed by their tenure as renters or owners. The first step is to assess, via a pre-determined happiness threshold, whether they are satisfied in their current location. If they are not satisfied, then they search the available (vacant) parcels up to 15 times (approximating a scan of the real estate section) to assess alternatives. If the utility of the alternative parcels exceeds the current location utility, they become part of a consideration set. After the consideration set is complete, the location with the highest neighborhood utility is chosen.

Note that the distance effect in choosing a social network, and the social network effect in choosing a neighborhood, constitute the recursive relationship between neighborhoods and networks as hypothesized in Figure 1 earlier. These effects may be turned off simply by setting the parameter weights to zero, to test alternative model structures. Effects may also be adjusted by changing the level of aggregation (e.g., block and blockgroup effects).

When a household moves to a new location, it induces a state change in the old and new parcels. First, the old parcel becomes vacant. The new parcel will be occupied, but whether it will be a rental or owner-occupied unit is not yet clear. To enable transitions between the rent and own state, both parcels and households keep track of their prior rent or own state.

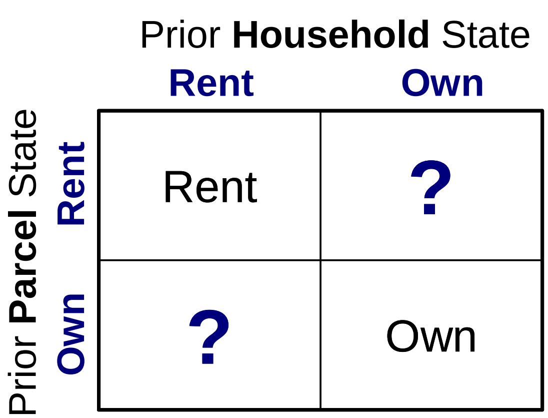

Figure 27. Transitions Between Rent and Own Status

As illustrated in Figure 27, if both the parcel and the household were formerly of "rent" status, that status is retained. The same logic is applied to the "own" status. But if the prior parcel state was "rent" and the household was formerly "own", or vice versa, the new state is uncertain. A simple way to handle this uncertainty is to flip a coin for determining the new state under these conditions. Despite this simple assumption, tracking ownership status and enabling transitions between "own" and "rent" status is important to the model formulation. This assumption enables more adequate calibration relative to the ownership migration patterns introduced in Chapter 3.

As mentioned, the dynamics of agent-based decisions are implemented using asynchronous evaluation. Such evaluations occur at a frequency of times derived from an exponential distribution of owner or renter move rates. The exponential distribution invokes a Poisson process and in the aggregate appoximates a first-order delay in the stock and flow representation of system dynamics.

4.4 Simulating Spatial Disparity

The dynamics of household choice described above create a system of migration patterns that may be calibrated relative to the observed migration patterns, as described in the next chapter. Although this project has not yet produced a calibrated solution, in this section I outline how a measure of spatial disparity was developed to examine the dynamics of a single simulation.

While a number of measures have been proposed to assess spatial disparity (i.e., Chakravorty 1996), I utilized the widely known Gini coefficient (Gini 1921) applied to incomes averaged at the blockgroup level. The Gini coefficient is evaluated as a fraction between zero and one, with higher values representing greater inequity. The spatial dimension of this inequity is provided by comparison among the 28 blockgroups containing households. Equation 3 describes the formula used to assess the Gini coefficient for the population of 28 blockgroups in this model.

Equation 3. Formula for Measuring Gini Coefficient

$$G = \frac{1}{n} \cdot \left(n + 1 - 2 \cdot \frac{\sum_{i=1}^{n} (n+1-i) \cdot y_i}{\sum_{i=1}^{n} y_i} \right)$$

where

G	= Gini coefficient
n	= number of blockgroups (28)
i	= index of blockgroup in non-decreasing rank order
y_i	= average income of i^{th} blockgroup

Using the Gini coefficient by blockgroup, Figure 28 reveals the dynamics of spatial disparity for a parameter set resulting from the optimization described in the next chapter. From a starting point of ~ 12%, spatial disparity increases to amost 16% over the first two years (the calibration period for migration patterns) but eventually declines to the starting level, with some variation.

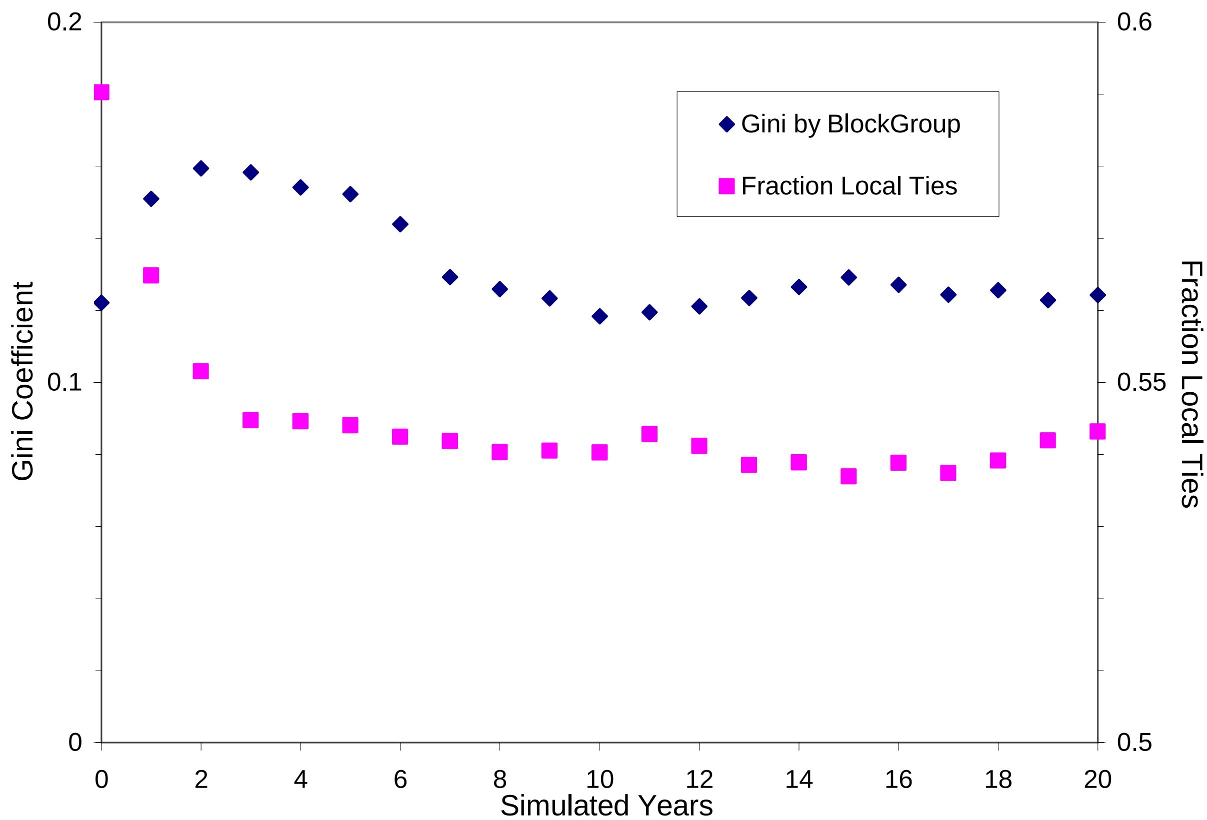

Figure 28. Gini Coefficient and Fraction of Local Ties over Years of a Single Simulation

The secondary (right-most) vertical axis of Figure 28 reveals the dynamics of the social network, measured as the fraction of each household's local ties within their home blockgroup. The fraction of local ties declines from the initial value of ~60% to around 54% over the years of simulation. The dynamics of the social network exhibit a first order effect as moves are generated and some ties are broken. Indeed, in this simulation some households were completely disconnected from their social network. In contrast, the dynamics of spatial disparity exhibit a lagged effect, increasing with the migration patterns but ultimately declining.

These dynamic simulation results provide an example of how a calibrated result may be investigated. Extreme caution must be taken in interpreting these results. The importance of calibration is described in the next chapter, along with the range of outcomes that fit reasonably to the observed migration patterns. However, even a "good" solution should be subject to skepticism until it has been validated dynamically over a time frame of interest to policy makers. The simulation shown in Figure 28 has not been adequately validated, as the calibration results are as yet inconclusive. Nonetheless, insights have been gleaned from the model development and calibration process, as are documented in the pages that follow.

CHAPTER 5. CALIBRATION RESULTS

This chapter highlights the crux of this research in articulating and exploring a calibration method for the agent-based model articulated in the previous chapter. Because of the inherent randomness in the model, all results presented in this chapter are averaged across batches of 25 simulation runs per experiment. The sections that follow first describe feasible model alternatives that may be tested within the articulated simulation framework. Then the full calibration strategy is described, followed by results that have been obtained toward this end.

5.1 Specifying Alternatives

As described in the previous chapter, the spatial dynamic behavior of this model stems from the two choices made by household agents: 1) the choice of friendships (social networks), and 2) the choice of where to live (neighborhoods). Each of these choices uses a utility formulation of weighted effects, where the weights are adjusted during model calibration. Different combinations of effects create alternative model structures.

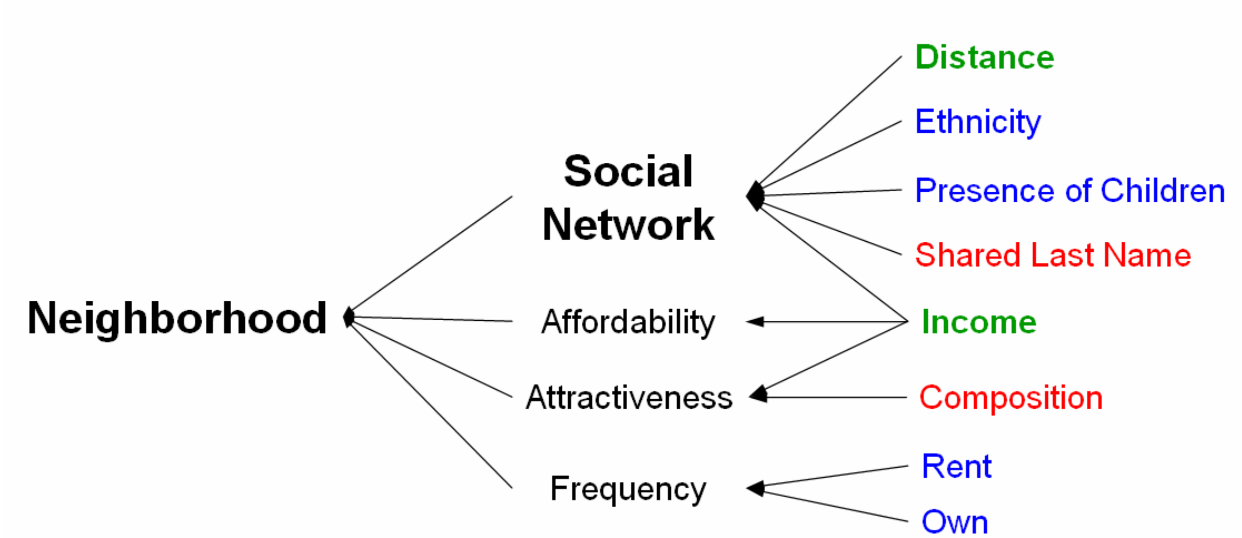

Figure 29. Model Structure Alternatives

The distance and income effects in Figure 29 are central to the full model structure as well as the prototype model. These effects are structured so that the greater the Euclidean distance between parcels, and the greater the income difference between households, the less likely the households are to have a social network connection. Neighborhood choice depends in part on the social network effect, expressed as the fraction of a household's connections within the prospective blockgroup area. The distance effect in choosing a social network, and the social network effect in choosing a neighborhood, constitute the recursive relationship between neighborhoods and networks as hypothesized in Chapter 1.

In addition to distance and income, several other effects were included in the full data-driven model. The rent or own status, informed by the Census, is used to vary the frequency of evaluating a move. Census-informed ethnicity (in the form of a "white" or "nonwhite" racial dichotomy), and presence of children were also included as factors in social network choice for the full model. Feasible model extensions could include shared last name aand housing composition.

Identifying which parameters are worth testing is a model boundary choice informed in part by the data analysis in Chapter 3. For example, friendships among children were revealed by interviews to be a significant factor in neighborhood networks. And shared last name could be included as a proxy for family connections. Income has already been included as an effect in social network choice as well as neighborhood choice, through the affordability and attractiveness terms. Another effect on neighborhood attractiveness could derive from its housing composition, such as presence or persistence of vacant units in the neighborhood. The real-life effect of excess vacancy in some Danville neighborhoods is to enable spaces for illegal activities such as drug dealing, meth labs, and vandalism. This, in addition to aesthetic concerns, could make vacancy a negative effect. Other composition effects could include ownership fractions and ethnicity.

Table 5 outlines which model parameters were selected for calibration. This analysis does not include the shared last name and housing composition effects postulated above.

Table 5. Description of Parameters Varied to Create Alternative Model Structures

Name	Description	Lower Bound	Upper Bound
locNorm	Normalizing Constant, Denominator of Parcel Location Utility	0.1	2
locConst	Linear Constant for Parcel Location Utility	0	1
locAfford	Weight on Household's Affordability Constraint	0	1
locAttract	Weight on Income-based Attractiveness of Block	0	1
locSN	Weight on Fraction of Social Network in Blockgroup	0	1
localProb	Probability of Forming Local Ties within Blockgroup	0.2	0.7
nwNorm	Normalizing Constant, Denominator of Social Network Utility	0.1	2
nwConst	Linear Constant for Household's Social Network Utility	0	1
nwChild	Weight on Boolean Effect of Children on Social Network Utility	0	1
nwDist	Weight on Effect of Block Distance on Social Network Utility	0	1
nwInc	Weight on Effect of Income Difference on Network Utility	0	1
nwRace	Weight on Boolean Effect of Racial Similarity on Network Utility	0	1

Combinations of these effects specify alternative model structures for calibration. Defining alternatives enables us to discern which combinations of effects match the observed migration patterns best. The next section outlines the strategy that was employed to conduct these tests.

5.2 Strategy

This section presents an overview of the calibration methodology employed in this project. The following are steps toward calibration of an agent-based model. These steps are not necessarily clear-cut and sequential. Grimm and Railsback (2005, pp. 343-345) outline the following calibration methodology for pattern-oriented modeling:

1. Using independent analysis, specify parameters that can be specified (e.g., location of parcels/owners, income from Census). Identify which parameters remain uncertain.
2. Specify ranges of values for uncertain parameters. (Note that steps 1 and 2 involve setting practical model boundaries.)
3. Create a set of permutations of parameter values across their possible ranges.
4. Define patterns to filter out unacceptable parameterizations. Here, the owner migration patterns described in Chapter 3 are used.
5. Design a scenario for the circumstances under which the filter patterns were observed. Here, the initial circumstances were set using 2001 owner data for the initialization, close to the 2000 census data for attribute specification.
6. Run simulations for all permutations of model parameterizations. Save output variables to match with patterns, as well as major model predictions.
7. Determine the parameter combinations (model structures) that reproduce the filter pattern at a coarse level.
8. Perform sensitivity testing on selected parameters within combinations, assessing influence in simulated output.
9. Assess structural robustness of model through testing of effects at the model boundaries testing.

Just as with the overall modeling process (as represented earlier in Figure 2 and Figure 5), the calibration process is also iterative. Nonetheless, these steps serve as a reasonable guide to making models more than abstract. Figure 30 illustrates the overall flow of the calibration process just described to create a model that makes sense in the real world.

At the top of Figure 30 are the initial steps. First, parameters are set where possible using data analysis – for income, race, parcel location, etc. Then broad ranges are initially chosen for the remaining uncertain parameters – whether they are negative or positive, for example. These bounds are readily adjusted using the OptQuest optimizer of the AnyLogic software. Alternative model structures are specified as combinations of parameters for the effects just described.

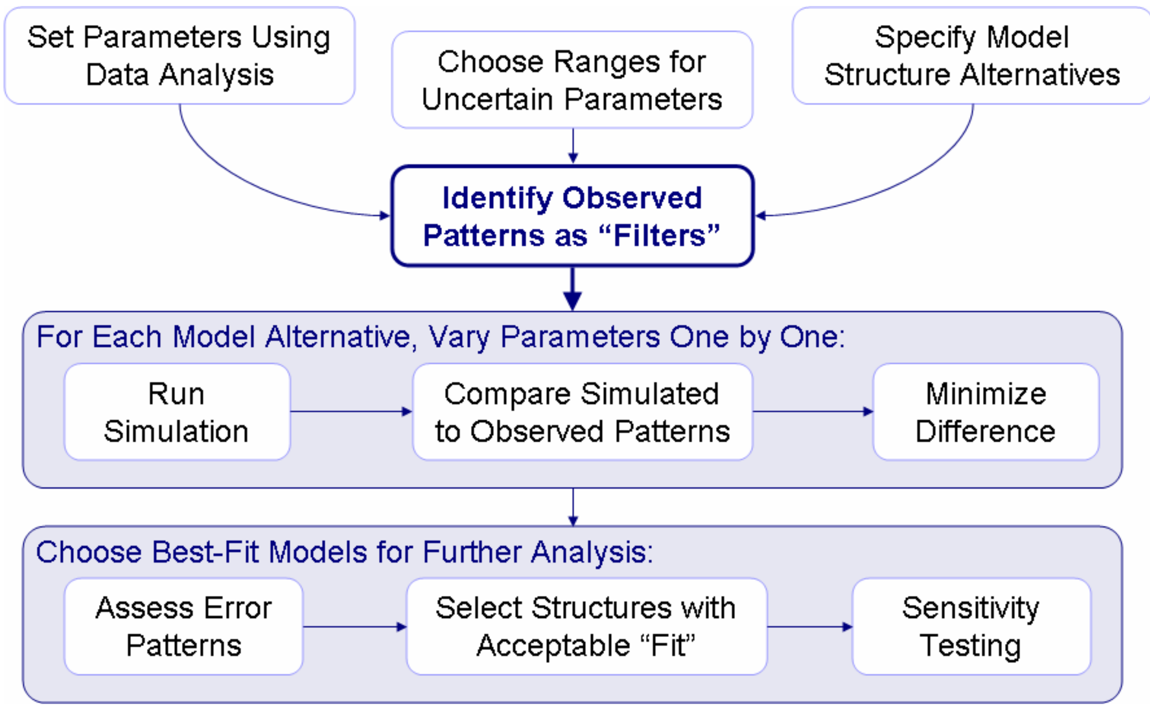

Figure 30. Calibration Strategy

The most critical step in this calibration strategy is to identify observed patterns that will serve as “filters” for acceptable models (Grimm and Railsback 2005). These patterns enable “decoding” of model structure (Wiegand et al. 2003). In this case, I use the owner data time series to create observed migration patterns.

With such patterns in place, each model alternative is tested by varying its constituent parameters one by one using the optimizer. The simulation is run, and simulated migration patterns are compared with the observed migration patterns. An error term is specified to capture the difference between simulated and observed patterns, with a penalty for excess parameters.

The goal is to minimize the error term. For this analysis, two error terms – total move errror and directional error accumulated by blockgroup – are combined in a single objective function.

Once each model alternative has minimized the error in its set of parameters, the best-fit models are selected for further analysis. Because the observed migration patterns include a subset of specific observations, the error patterns may be assessed to understand why some structures fit better than others. Structures with acceptable overall fit are selected to focus on further sensitivity testing. The focus here is on overall fit, rather than the statistical significance of individual parameters. In creating the agent-based algorithms, I have deduced from aggregate data and now try to reproduce broad patterns. In the event that no model fits acceptably, the error patterns may still be assessed for greater understanding.

5.3 Testing

The AnyLogic software comes with a built-in optimization package that enables a wide variety of experiments under uncertainty. Figure 31 shows a screenshot of the optimization window in AnyLogic. The left panel reveals the parameter names that have been selected to vary, along with the best value attained to date. Appendix G documents how the parameters are encoded for network and location choice, as described in the previous chapter.

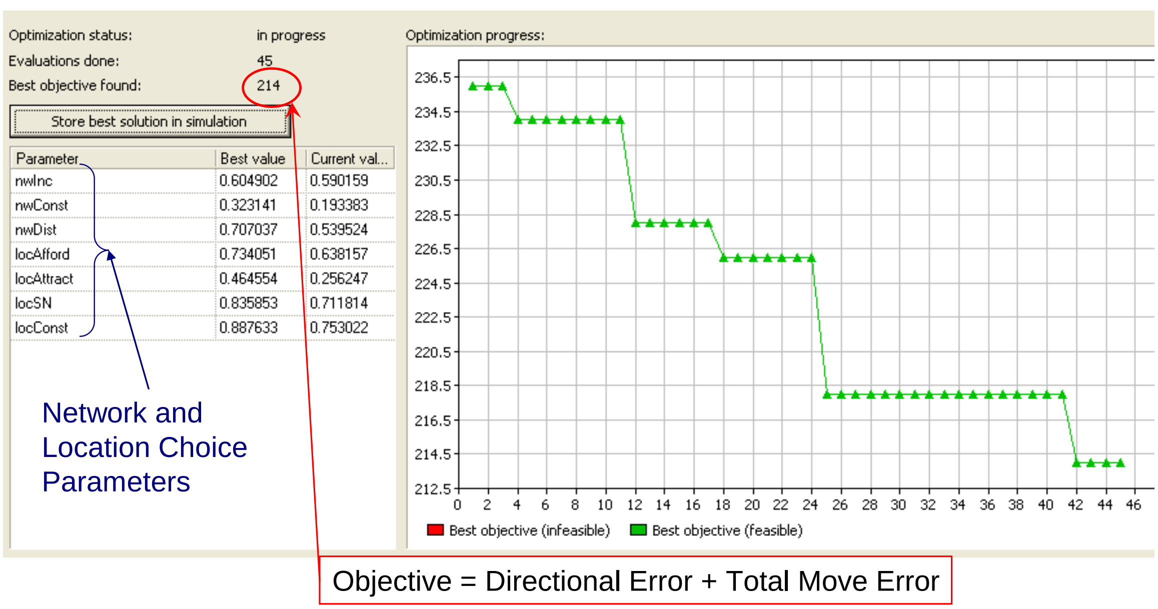

Figure 31. View of Optimization Process

Each experiment is run for the first two years and evaluated relative to the observed migration patterns. The objective function combines directional error (accumulated by blockgroup relative to the matrix of observed migration in Table 13 of Appendix E) and cumulative move error to ensure that approximately the right number of intra-community owner moves are made. If a new test does not produce better results than the previous test, the previous "best" is carried over in the right-most plot of Figure 31. Note that this example is illustrative only, as it reflects a single simulation outcome rather than a batch average as described below.

Each parameter is varied Monte Carlo style such that random combinations are tested across the range of possible parameter values. The bounds are defined in the optimization setup and are tested in this optimization process. Parameter weights and constants were bounded between zero and unity, since each effect was normalized. The scaling constants (in the denominator of the utility equations) were bounded between 0.1 and 2 to reflect a range of reasonable elasticity. And the probability of evaluating social connections within the home blockgroup, encoded as the "localProb" term in Appendix G, varied between 20% and 70%.

5.3.1 Solution Space

The single "best" outcome in the optimization process just described (e.g., see Figure 28) is less useful than a range of outcomes in understanding model behavior. The simulated errors were thus exported for further analysis, alongside each set of parameter values. The plot in Figure 32 shows the error produced for each of 1000 experiments, where each experiment is averaged across a batch of 25 simulation runs over a 2-year period.

The total move error exhibits a linear relationship with directional error by blockgroup. Because the total move error is considered as the absolute value of the difference between simulated and observed moves, this relationship appears in the shape of a V in Figure 32. As described above, the total and directional errors (accumulated by blockgroup) are added in a single objective function. This objective function is minimized in the "good" solution space highlighted, representing combined (total + directional) error of less than 230 moves. While an ideal situation would be to reach zero on both axes, such a goal has been unattainable in this set of experiments.

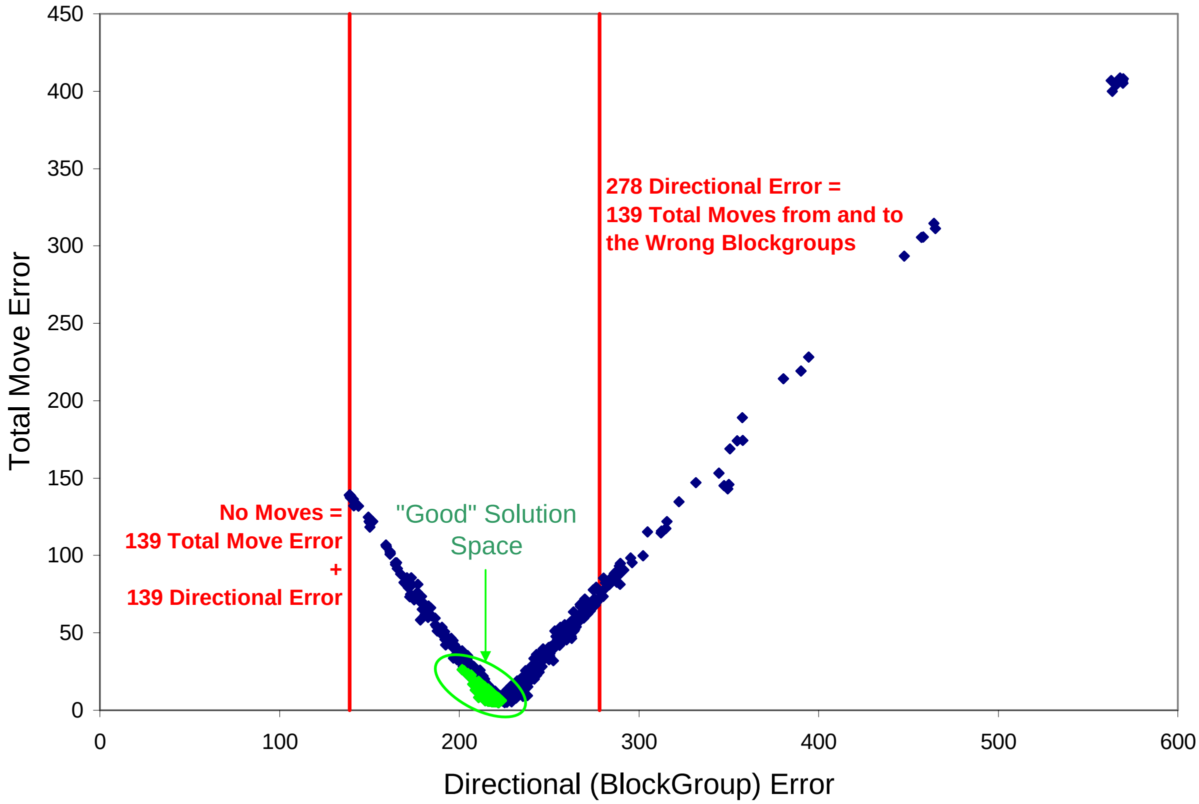

Figure 32. Total Move Error and Directional (BlockGroup) Error for 1000 Experiments

A range of outcomes is revealed in Figure 32, illustrating the relationship between total move error and directional error. If directional error were minimized completely, the trivial solution of no moves would be preferred over more meaningful solutions. This is because the observed 139 intra-community owner moves would produce equivalent errors in both the cumulative and directional dimensions. Total move error is minimized when directional error reveals approximately 220 moves. The difficulty of matching the directional dimension is largely due to the substantial number of zeros[17] in the observed migration patterns shown in Appendix E. Despite this difficulty, the "good" solutions shown in Figure 32 are significantly better than a directional error of 278 (the equivalent of matching total moves from and to all the wrong places).

[17] In the 28 x 28 owner move matrix (solely for internal migration patterns), there are 784 total cells. Only 86 of these are nonzero, containing the 139 total moves. This leaves 698 cells containing zero moves. Most of these cells are matched in the solution space.

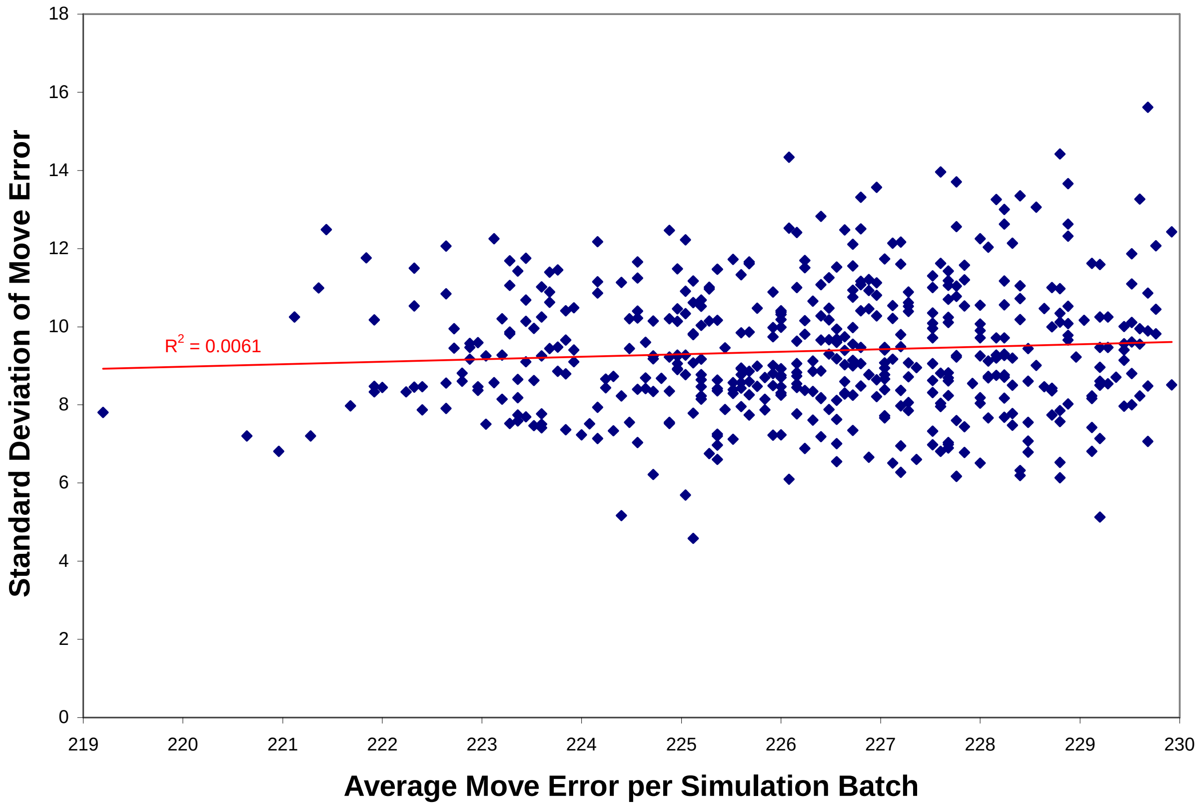

Figure 33. Variation of Move Error by Amount of Error Among Filtered Solutions

These 486 "good" solutions were selected by filtering for a combined error (objective function) of less than 230 directional and total moves. Figure 33 displays the variation of this error within each batch of simulation runs, revealing that the standard deviation of error is not significantly related to the amount of error per se (evidenced by an R-squared term of nearly zero for a linear relationship form). The standard deviation for this average error across the 25 runs in each experimental batch varied between 5 and 15 moves, averaging 10 moves of deviation from the average error.

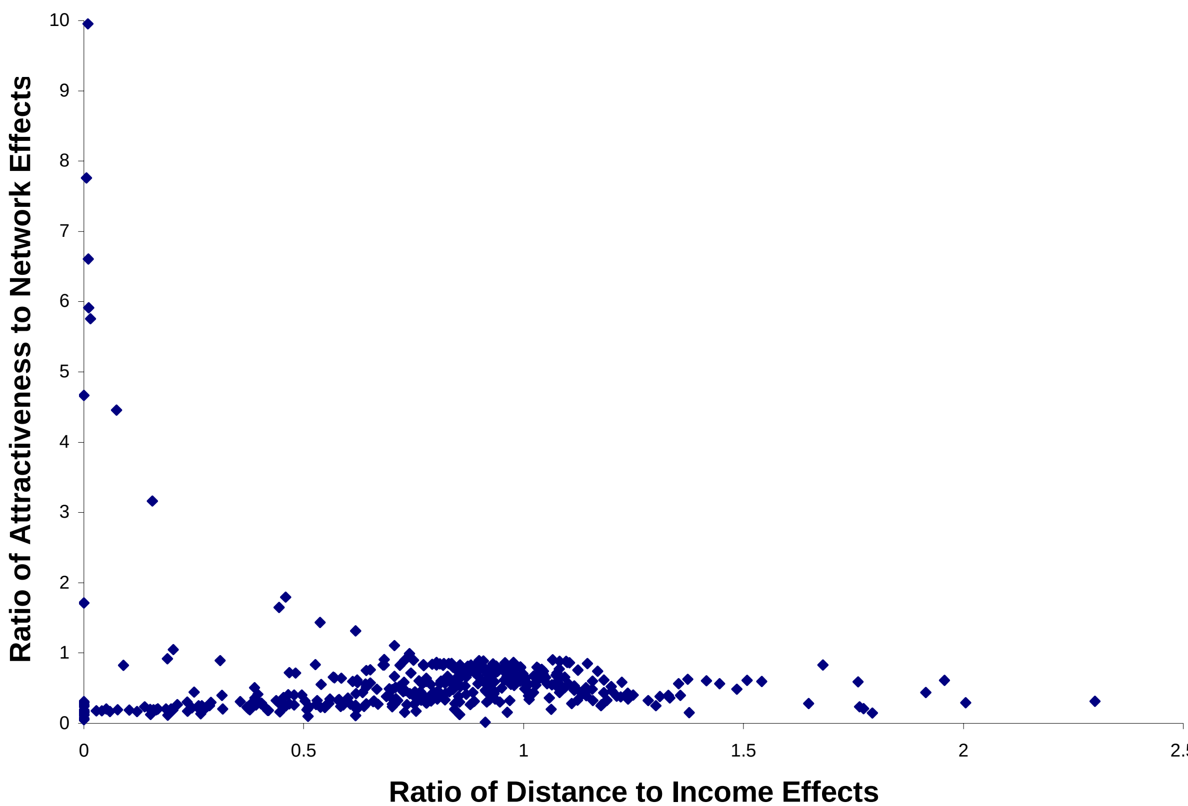

Figure 34. Filtered Patterns of Parameter Effect Ratios

The filtered set of solutions may now be used to investigate the multi-dimensional parameter space in search of insightful patterns. Figure 34 reduces these dimensions by comparing ratios of effects. The vertical axis shows the ratio of income-based attractiveness to social network effects on location choice, while the horizontal axis reveals the ratio of distance to income effects on social network choice. A handful of data points high on the vertical axis and low on the horizontal axis were excluded from this plot for visualization purposes.

Although the overall distribution envelope in Figure 34 appears to be somewhat exponential in nature, a significant result is that most of the ratios in this set of solutions are less than one. This implies that the attractiveness effect is generally weighted less than the social network effect on location choice, and the distance effect is generally weighted less than the income effect on social network choice. Note that a low parameter weight does not imply a less significant contribution to simulated migration patterns. Indeed, sensitivity testing of the income-based attractiveness effect revealed that its parameter weight has a very strong effect on inducing intra-community migration. Similarly, the social network effect is frequently zero, as an individual household has 6 social ties on average, spanning a possible space of 28 blockgroups.

Therefore, the higher weight on the social network effect does not necessarily imply a stronger impact on the resultant migration error.

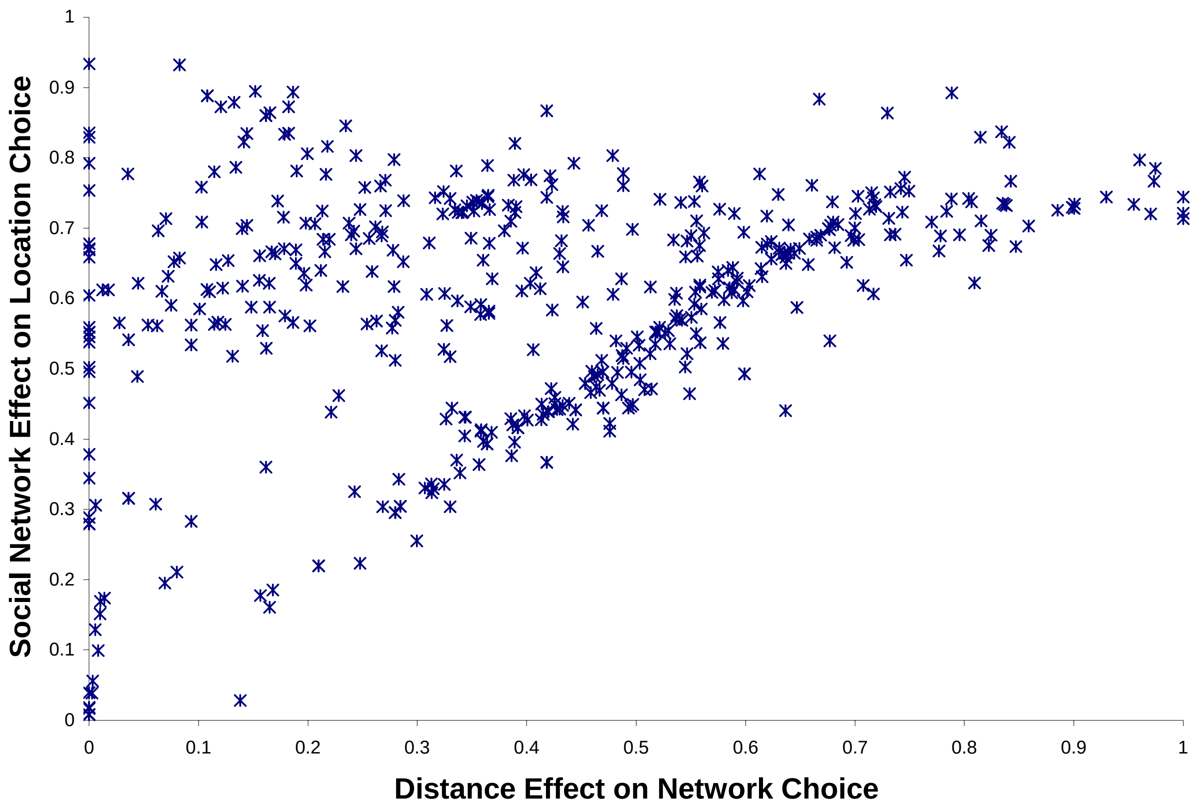

Figure 35. Filtered Comparison of Spatial and Social Influences on Choice

An examination of the recursive relationship between spatial and social effects in the solution space is provided by Figure 35. The vertical axis displays the weights for the social network effect on location choice, and the horizonal axis displays the weights for the distance effect on social network choice. While the relationship would not be readily reduced to a linear form, it appears to be a positive relationship. As the distance effect increases on network choice, so does the network effect on location choice. Also, when the distance effect is low, more variation is observed in the social network effect. Although the nature of the relationship is ill-defined in algebraic terms (e.g., a linear trendline), it is significant to note the positive correlation among the filtered solutions.

A cluster analysis of parameter weights in the set of 486 "good" solutions was performed to delve deeper into the simulated patterns of effects. Figure 36 reveals the results of this analysis as performed in SPSS as a two-step cluster classifcation.[18] Note that each parameter was multiplied by 100 to improve the scaling of statistical analyses performed in SPSS. In addition to the parameter weights shown, both directional and total move errors were included in the cluster analysis.

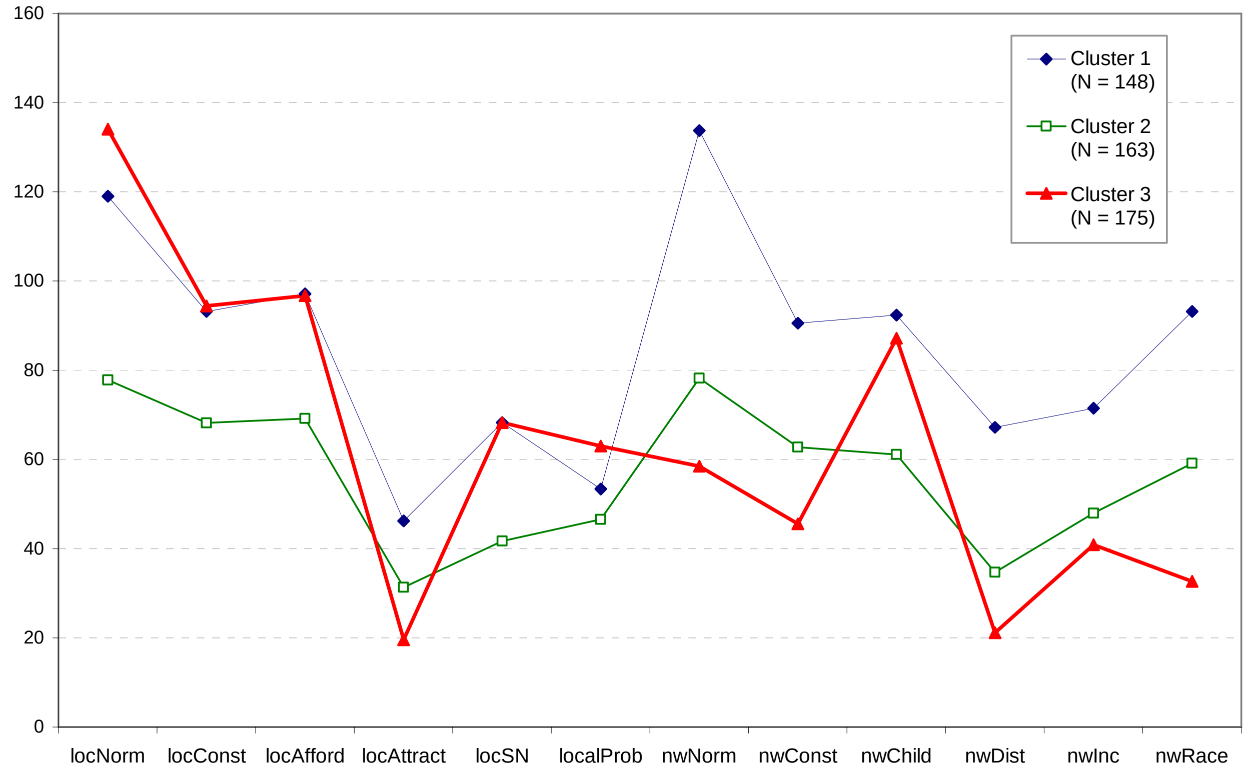

Figure 36. Cluster Analysis of Parameters in "Good" Solution Space

Three clusters of similar size resulted from this analysis. Figure 36 shows the means for each parameter weight in the cluster. The migration errors are not shown because they were so similar across the set, which had been filtered to have a cumulative error of less than 230. However, cluster 3 (triangles connected with a bold line) resulted in a slightly lower error in this analysis, with one less move error on average than the other two clusters. Cluster 3 overlaps with cluster 2 for several parameters, such as the affordability (*locAfford*) and social network (*locSN*) effects on location choice, as well as the constant term (*locConst*) for that choice. Moreover, the ratios explored in Figure 34 are corroborated by the average cluster results in Figure 36 (i.e., *locAttract/locSN* < 1 and *nwDist/nwInc* < 1). The probability of selecting social ties within one's

[18] The SPSS clustering criterion was BIC (Bayesian Information criterion) with a log-likelihood distance measure.

blockgroup (*localProb*) is above 50% for all three clusters. Although the distance effect in social network choice is more widely variable, this pre-filtering of choice within blockgroup boundaries is consistently substantial.

The filtered set of 486 solutions just examined has reduced the solution space by more than half. But this (still large) number of solutions precludes the appearance of a definitive answer. As noted in Chapter 2, such a definitive solution would not be expected for this kind of a "wicked" problem (Rittel and Webber 1973). It is substantial already that this research has produced a functional model formulation – not a definitive formulation, but rather one that may be revisited after adequate testing and analysis.

5.3.2 Regression Results

Building on the exploratory data analysis just presented, this section provides a summary of the regression results for the simulated output. Here, the independent variables were the parameter values for each of 1000 experiments performed. Each experiment consists of 25 simulation runs to compensate for random variations between runs. The dependent variable was the average move error, combining both directional and cumulative moves across each simulation batch.

Table 6. Regression Summary of Parameter Effects on Average Move Error

	Full Model	**Linear Model**
R-squared	.701	.478
Standard Error of the Estimate	50.6	66.3
Social Network Effect (*locSN*)	significant	NOT significant
Effect of Children on Network (*nwChild*)	NOT significant	more significant
Effect of Distance on Network (*nwDist*)	NOT significant	NOT significant
Probability of Local Ties (*localProb*)	significant	significant
Affordability Constraint (*locAfford*)	significant	NOT significant
Network Constant (*nwConst*)	significant	NOT significant

Table 6 provides an overview of the regression analyses performed on the simulation data. Two models were created for comparison of parameter significance. The full model includes nonlinear (squared) as well as linear forms of the 12 independent variables.[19] While the

[19] Each parameter was multiplied by 100 for scaling purposes prior to regression and transformation into squared terms.

full model has a better fit (i.e., a higher R-squared term) than the linear model, it also has twice as many independent variables.

Aside from the comparison of R-squared and standard error terms, Table 6 provides a qualitative comparison of the two models, highlighting which parameters appeared as significant. The significance of some effects depends upon the model. For example, the social network effect on location choice was deemed significant when squared terms were included, but not in the linear model. The effect of children on social network formation was not significant in the full model, but was more significant in the linear model. The distance effect on network formation appeared insignificant in both model forms, but the probability of local ties (a means of incorporating spatial effects within a discrete blockgroup area rather than along a spectrum of block distance) was significant in both. Regression details for each model are provided in Table 7 (linear model) and Table 8 (full model).

Table 7. Regression of Model Parameters (Linear Form) on Average Move Error

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	371.219	11.747		31.601	.000
	nwNorm	.577	.105	.315	5.524	.000
	nwConst	-.010	.203	-.003	-.051	.960
	nwInc	-.766	.158	-.214	-4.853	.000
	nwDist	-.196	.231	-.061	-.848	.397
	nwRace	-.629	.155	-.245	-4.048	.000
	nwChild	-.305	.181	-.085	-1.687	.092
	locNorm	.153	.104	.076	1.462	.144
	locConst	-1.218	.189	-.337	-6.435	.000
	locAfford	-.129	.143	-.045	-.897	.370
	locAttract	1.526	.156	.465	9.808	.000
	locSN	.237	.187	.066	1.269	.205
	localProb	-.762	.269	-.112	-2.831	.005

a. Dependent Variable: avgErr

The regression results in Table 7 reveal the coefficients and significance levels of the 12 linear parameters (and a constant term). The results appear straightforward in this form, though the linear model fit to the average error is not as close as the full model. As mentioned in the context of Table 6, several of these parameters have different significant levels from their nonlinear form in the full model.

Table 8. Regression of Model Parameters (with Squared Terms) on Average Move Error

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	558.212	25.550		21.847	.000
	nwNorm	.722	.300	.394	2.410	.016
	nwNorm2	-.002	.001	-.256	-1.677	.094
	nwConst	-1.907	.497	-.550	-3.833	.000
	nwConst2	.022	.004	.751	5.024	.000
	nwInc	2.828	.486	.791	5.815	.000
	nwInc2	-.034	.004	-1.115	-7.626	.000
	nwDist	-.265	.483	-.082	-.549	.583
	nwDist2	.002	.004	.070	.543	.587
	nwRace	1.149	.394	.448	2.918	.004
	nwRace2	-.014	.003	-.594	-4.298	.000
	nwChild	-.339	.543	-.094	-.624	.533
	nwChild2	-.003	.005	-.099	-.632	.528
	locNorm	.074	.277	.037	.266	.790
	locNorm2	.000	.001	.028	.223	.824
	locConst	-5.419	.555	-1.500	-9.758	.000
	locConst2	.035	.005	1.229	7.580	.000
	locAfford	-2.007	.406	-.710	-4.938	.000
	locAfford2	.015	.003	.610	4.435	.000
	locAttract	-1.955	.362	-.595	-5.397	.000
	locAttract2	.039	.004	1.224	10.648	.000
	locSN	1.934	.489	.537	3.956	.000
	locSN2	-.014	.005	-.366	-2.996	.003
	localProb	-5.297	1.419	-.782	-3.732	.000
	localProb2	.045	.015	.656	3.005	.003

a. Dependent Variable: avgErr

Table 8 lists the coefficients and significance of 24 parameters (plus a constant) for the regression of a model with both linear and squared terms. As noted in Table 6, this model fits the simulation error better than the linear model form, but does so at the expense of extra parameters. The significance (as measured by the t-value magnitude greater than 2, and p-value less than .05) tests reveal that both linear and squared parameters share significance levels. That is, both linear and squared income effects on network choice are significant. Similarly, both linear and squared distance effects on network choice are insignificant. Several of these parameter pairs also changed coefficient polarity between the linear and squared forms.

The results of this analysis are mixed. Social network influence on location choice is considered significant in a nonlinear form but not in the linear model. Distance effects on network choice are not considered significant, though the creation of local ties is significant on reducing migration error. Spatial and social effects are interlinked by migration patterns, and are suggestive (if not conclusive) of significance in alternative model forms.

CHAPTER 6. CONCLUSIONS

It is a bittersweet feeling to come to the end of a project and realize it is only the beginning. On the one hand, the results described in the previous chapter do not conclusively demonstrate the effect of social networks on spatial disparity, or even migration patterns. On the other hand, the results are suggestive enough to warrant further exploration using the modeling methodology demonstrated in this project. Such exploration is warranted in two ways: (1) further analysis of filtered solutions and their implications for the social link with spatial disparity; (2) revisiting the problem and beginning another cycle of modeling as suggested by both Figure 2 and Figure 5 earlier.

This research represents a significant step toward addressing the "wicked" problem (Rittel and Webber 1973) of spatial disparity using agent-based modeling. As with other large models of rich spatial and temporal detail (e.g., BenDor and Metcalf 2006), the model generated in this project may be considered a sort of knowledge repository to consolidate alternative forms of information relevant to the Danville case. The inclusion of household choice in a spatial dynamic simulation model crosses a pragmatic frontier in representing multidimensional complex systems (Agarwal et al. 2002). More broadly, the structural template of the model may be applied to other metropolitan areas facing similar dilemmas of disparity.

As articulated in Chapter 1, the goal of this study was to examine how social networks can shape community structure through neighborhood choice. The relationships illustrated in Figure 1 serve as a dynamic hypothesis to guide model conceptualization for the case of Danville, IL, USA. A variety of data sources provided a rich starting template of household agent attributes in a spatial dynamic simulation model. Spatial owner migration patterns accumulated at the Census blockgroup level serve to inform model development and calibration.

A prototype model (built without quantitative data) accelerated the model development cycle by testing heuristics in a two-neighborhood system. The full model (see CaseTown Package and Model Structure in Appendices F and G) was built with a combination of java classes and AnyLogic active objects. Spatial data at a variety of scales (e.g, parcel, block, blockgroup) were processed using a set of algorithms to structure a set of state objects that provide context for the household agent objects. Household attributes were inferred probabilistically from Census 2001 distributions in conjunction with available owner-occupied household locations for 2001. The initial social network among households is constructed from the same network choices that are made over time.

Network and neighborhood choices are formulated such that the utility of a decision depends upon parameter weights that may be estimated indirectly from observed migration patterns. The significant amount of randomness in this model can produce divergent, path-dependent outcomes for the same parameter settings. While the random aspects of initial network formation may be fixed using a random seed, the dynamics of the full model have not been fixed in this way. AnyLogic software experts are presently helping to excavate the source of this trouble and determine its implications.

6.1 Insights

Balanced spatial preferences in the prototype model led to a stronger emphasis on income in determining choices and thus disparity between neighborhoods. Those who can afford to move to the more attractive neighborhood do so, producing an agglomeration of higher income individuals in the attractive neighborhood. At the same time, balanced preferences result in a higher cohesion level, so that spatial disparity and fragmentation do not coexist at their extremes under this model formulation.

For the applied CaseTown model, the results of 1000 simulation experiments provide insight into the effects of alternative model structures on recreating migration patterns. Each experiment allows for variation across simulation runs by averaging directional and cumulative move error over a batch of 25 runs. The solution space was more than halved by filtering results below a combined error (objective function) of 230 moves. The resulting 486 alternatives were examined for patterns in the parameter space linking social and spatial effects. Linear regression was performed on the set of 1000 experiments to test the significance of the choice parameters in predicting migration error.

The significance of social network influence on neighborhood choice appears in a nonlinear regression model but not in a linear form. The distance effect on network choice is not significant, but the probability of forming local ties is a significant effect generally estimated to be greater than 50% likelihood in the home blockgroup.

The analysis stage of the modeling process (e.g., Figure 5) warrants at least as much time and effort as the model formulation, though it is often cut short by project deadlines. Fortunately, the data-intensiveness of this model necessitated analysis throughout the research process, so that insights were archived along the way. For example, the measurable significance of shared last name on spatial proximity suggested the influence of social ties on spatial patterns.

The owner-occupied household data over a two-year time series revealed a northwest migration pattern within Danville, as congruent with anecdotal evidence and expectations. This internal migration pattern contrasted with the overall flows of outmigration and new ownership. Outmigration from Danville was evident in the more affluent neighborhoods, ostensibly where property values could be recouped as mobile professionals relocated for better job opportunities. An influx of ownership in lower income areas (such as the neighborhood of focus for this project) revealed the significance of renter to owner transitions. Since the property values had reached bottom in lower income areas, landlords were willing to sell houses to tenants on contract after they made enough money from renting.

Because the renter-owner transitions are significant, and the renter-owner ratio is considered to be an indicator of neighborhood stability, renters and owners are modeled explicitly in this project. While an influx of renters could bring instability to a neighborhood, the transition to ownership status enables an avenue for extending tenure and strengthening social networks (as suggested by the causal map in Figure 26). Although a flexible framework has been created, further analysis of renter/owner dynamics is warranted.

This research has developed a model and theoretical framework to examine the dynamics of spatial and social influences on neighborhood and network choice. While the results are mixed, they illustrate how a computer model may be used to shed insight on the spatial dynamics of emergent social phenomena such as spatial disparity. I utilize fieldwork that elicits perceptions of place to help structure the model and also reveal its limitations. With further developments, the model may be used as a computational laboratory to explore the effects of alternative policy settings, individual assumptions and institutional structures.

6.2 Model Extensions

The most immediate and direct extensions of this work would be to extend the analysis of simulated migration patterns. To do so, the error between simulated and observed moves may be structured in an objective function such that cells in the migration matrix are weighted by the number of moves therein. The majority of zero cells would then matter less to the objective function than cells with significant flows of homeowners. Refinement of this error would enable better examination of parameter effects.

While the mechanics of migration have been emphasized in this iteration of the modeling cycle, they must still be linked to measures of spatial disparity. Such disparity could be measured in different ways. The prototype model assessed the gap in average income between two neighborhoods as an indicator of disparity. This income gap was compared with network characteristics such as the fraction of social ties crossing neighborhood boundaries. While the measure of local social ties has been employed in the full model across 28 blockgroups, a single measure of spatial disparity has not been articulated at the blockgroup scale.

Extensions of model scope could include familial relationships derived from shared last name. Spatial dimensions of family ties are measurable from the distance analysis in Chapter 3. Similarly, the presence of children was more significant than factors such as race and block distance in influencing social network choice. As noted in the previous chapter, model extensions focusing on neighborhood choice could focus on the aggregate feedback of housing composition on prospective household choice.

Negative (balancing) effects from excess vacancy and rentals, as well as mixed effects from preferences for similarity or diversity of race or ownership status could be incorporated. Indeed, the abstract segregation model introduced by Schelling (1971, 1978) suggests that a preference for diversity is required to overcome even a slight self-sorting tendency. Well-designed mixed-use neighborhoods could help to promote such individual-level preferences.

6.3 Policy Analyses

Once a feasible set of solutions has been derived from sensitivity analysis, more specific experiments may be conducted to test scenarios of migration shocks. Controlled variations of renter-owner transitions and evaluation frequencies would also be warranted. These experiments require resolution of the simulated random seed problem to ensure reproducible results. With these adjustments, dynamics of spatial disparity indicators may be juxtaposed with community network structures in a coherent manner.

The intrinsic policies that may be tested with this model include whether it is a good idea to allocate resources toward neighborhood association formation. Specifically, policy analysis could aid in addressing the following questions:

- ❖ How can local networks develop at a fine enough scale of neighborhood to create a coherent cluster? A dense spatial cluster of social networks could become myopic and insular, or it could become an organization empowered to transact directly with the city. A neighborhood association could induce the latter.
- ❖ How can a collective sense of place be incubated in such a network to balance rapid migration dynamics? A policy encouraging renters to transition to homeowner status could help to stabilize neighborhoods. Alternatively, such a policy may not matter much beyond a particular tipping point of neighborhood composition (be it along dimensions of income, tenure, vacancy or race).
- ❖ How does spatial disparity relate to social fragmentation? This question requires explicit formulation of both disparity and fragmentation, despite their conceptual linkage the in the introduction to this research. Counter to expectations, the prototype model measures for disparity and fragmentation did not coexist at the same instance in spacetime.

6.4 Implications

In the end, the purpose of this project has been to examine the feasibility of this modeling and calibration methodology as applied to the persistent problem of spatial disparity. While this is only a modest step toward understanding the problem, it is a more groundbreaking step in terms of suggesting alternative ways of examining the problem.

The simulation model presented herein provides the opportunity to test understanding of migration dynamics using internally consistent assumptions. In this way, the model framework serves as a repository for knowledge arriving in various forms, from qualitative explanations to quantitative data. Alternative model structures enable multiple valid perspectives on the problem, and the real-world case provides a sort of “ground truth” to *humble* the model and reveal its limitations. The goal of modeling is to generate iterative learning about the problem. The benefits of this iteration are not restricted to modelers – it engenders dialogue between policy stakeholders and residents. One of my interviewees reminded me that my own participation was influencing the growth of the neighborhood association, for example.

When we use the power of direct experience, we can creatively think of how our own decision rules operate, and learn to translate these rules into algorithms. While we humans have a great advantage in deeply understanding our human subjects from within, this renders us our own harshest critics about the validity of any assumptions at all. After all, humans are unpredictable! As soon as we attempt to model ourselves, we have transcended the selves we wished to represent.

It is because of the impossibility of ever modeling ourselves completely that we must always iterate in the modeling process. No answer will be conclusive, but we can take what we learn and apply it to further questions, inquiring into a universe of "whys" as we get closer to understanding our fundamental nature.

The pattern-oriented modeling process applied herein uses patterns to decode information otherwise unavailable to the scientist (Wiegand et al. 2003, Grimm and Railsback 2005, Grimm et al. 2005). In this research, migration patterns have enabled examination of possible recursive relationships between social network and neighborhood choice. The results are neither conclusive nor inconclusive, but rather suggestive of further research in this domain.

The process of learning by experimenting enables deeper understanding of the problem at hand. Despite an apparent quest for realism in modeling, it is this process of exploration that enables intuition to build from the experiment of modeling ourselves. In addressing the problem of spatial disparity, I have focused on insights rather than specific answers. Understanding the model structure is central to understanding model behavior. Because so much uncertainty exists in many of the assumed parameter values, much insight can be gleaned from observing patterns of simulated and empirical behavior over spacetime.

REFERENCES

Adams, P. 1995. "A Reconsideration of Personal Boundaries in Space-Time." *Annals of the Association of American Geographers* 85(2): 267-285.

Adams, P. 1998. "Network Topologies and Virtual Place." *Annals of the Association of American Geographers* 88(1): 88-106.

Agar, M. 2004. "We Have Met the Other and We're All Nonlinear: Ethnography as a Nonlinear Dynamic System." *Complexity* 10(2): 16-24.

Agarwal, C., G. M. Green, J. M. Grove, T. P. Evans and C. M. Schweik. 2002. *A Review and Assessment of Land-Use Change Models: Dynamics of Space, Time, and Human Choice.* General Technical Report NE-297. Newton Square, PA: U.S.Department of Agriculture, Forest Service, Northeastern Research Station. http://www.fs.fed.us/ne/newtown_square/publications/technical_reports/pdfs/2002/gtrne297.pdf

Aitken, S. C. 1990. "Local Evaluations of Neighborhood Change." *Annals of the Association of American Geographers* 80(2): 247-267.

Arthur, W. B. 1988. "Urban Systems and Historical Path Dependence." In *Cities and their Vital Systems: Infrastructure, Past, Present, and Future*, ed. J. H. Ausubel and R. Herman. Washington, DC: National Academy Press, 85-97.

Atkin, R. 1974a. "An Approach to Structure in Architectural and Urban Design. 1. Introduction and Mathematical Theory." *Environment and Planning B* 1: 56-67.

Atkin, R. 1974b. "An Approach to Structure in Architectural and Urban Design. 2. Algebraic Representation and Local Structure." *Environment and Planning B* 1: 173-191.

Axelrod, R. 1997. *The Complexity of Cooperation*. Princeton, NJ: Princeton University Press.

Barabasi, L. A. 2002. *Linked: The New Science of Networks*. Cambridge, MA: Perseus.

Barabasi, L. A. and R. Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 386(5439): 509-512.

Batty, M. 2005. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based Models, and Fractals*. Cambridge, MA: MIT Press.

Ben-Akiva, M. and S. R. Lerman. 1985. *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

BenDor, T. K. and S. S. Metcalf. 2006. "The Spatial Dynamics of Invasive Species Spread." *System Dynamics Review* 22(1): 27-50.

Borshchev, A. and T. Popkov. 2006. "Getting Started with AnyLogic Software and Agent-Based Modeling." In *Proceedings of the 24th International Conference of the System Dynamics Society*. Nijmegen, The Netherlands: System Dynamics Society. http://www.systemdynamics.org/conferences/2006/proceed/papers/BORSH450.pdf

Bowles, S. 2001. "Individual Interactions, Group Conflicts, and the Evolution of Preferences." In *Social Dynamics*, ed. S. N. Durlauf and H. P. Young. Washington, DC: Brookings Institution Press, 155-190.

Brown, D. G. and D. T. Robinson. 2006. "Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl." *Ecology and Scoiety* 11(1): 46.

Buchanan, M. 2002. *Nexus: Small Worlds and the Groundbreaking Science of Networks*. New York, NY: Brookings Institution Press.

Burt, R. S. 1987. "Social Contagion and Innovation: Cohesion versus Structural Equivalence." *American Journal of Sociology* 92(6): 1287-1335.

Chakravorty, S. 1996. "A Measurement of Spatial Disparity: The Case of Income Inequality." *Urban Studies* 33(9): 1671-1686.

Conley, T. G. and G. Topa. 2002. "Socio-Economic Distance and Spatial Patterns in Unemployment." *Journal of Applied Econometrics* 17: 303-327.

Conley, T. G. and G. Topa. 2003. "Identification of Local Interaction Models with Imperfect Location Data." *Journal of Applied Econometrics* 18(5): 605-618.

Couclelis, H. 1986. "A Theoretical Framework for Alternative Models of Spatial Decision and Choice." *Annals of the Association of American Geographers* 76(1): 95-113.

Crane, J. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology* 96(5): 1226-1259.

Davis, J. A., T. W. Smith and P. V. Marsden. 2005. *General Social Surveys 1972-2004. ICPSR04295-v2.* Chicago, IL: National Opinion Research Center.

Degenne, A. and M. Forse. 1999. "Dynamics." In *Introducing Social Setworks*, London, UK: Sage Publications, 159-184.

Desbarats, J. 1983. "Spatial Choice and Constraints on Behavior." *Annals of the Association of American Geographers* 73(3): 340-357.

Dibble, C. and P. G. Feldman. 2004. "The GeoGraph 3D Computational Laboratory: Network Terrain Landscapes for RePast." *Journal of Artificial Societies and Social Simulation* 7(1). http://jasss.soc.surrey.ac.uk/7/1/7.html

Downs, A. 1999. "Some Realities about Sprawl and Decline." *Housing Policy Debate* 10(4): 955-973.

Downs, A. 2000. "Relating Sprawl to Urban Decline: Effects of Suburban Sprawl on the Decline of Metropolitan Areas." In *Costs of Sprawl 2000: TCRP Report 74*, ed. R. W. Burchell, G. Lowenstein, W. R. Dolphin and C. C. Galley. Washington, DC: Transportation Research Board, 373-440.

Durlauf, S. N. and H. P. Young, Eds. 2001. *Social Dynamics*. Washington, DC: Brookings Institution Press.

Epstein, J. M. and R. Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, DC: Brookings Institution Press.

Erdos, P. and A. Renyi. 1960. "On the Evolution of Random Graphs." *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5: 17-61.

Forrester, J. W. 1996. "System Dynamics and K-12 Teachers." D-memo D-4665-4. Cambridge, MA: Massachusetts Institute of Technology. 35.

Gilbert, M. 1998. "Race, Space, and Power: The Survival Strategies of Working Poor Women." *Annals of the Association of American Geographers* 88(4): 595-621.

Gimblett, H. R., Ed. 2002. *Integrating Geographic Information Systems and Agent-Based Modeling Techniques for Simulating Social and Ecological Processes*. New York, NY: Oxford University Press.

Gini, C. 1921. "Measurement of Inequality and Incomes." *The Economic Journal* 31: 124-126.

Gould, P. 1981. "A Structural Language of Relations." In *Future Trends in Geomathematics*, ed. R. Craig and M. Labovitz. London, UK: Pion, 281-312.

Granovetter, M. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6): 1420-1443.

Grimm, V. and S. F. Railsback. 2005. *Individual-based Modeling and Ecology*. Princeton, NJ: Princeton University Press.

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand and D. L. DeAngelis. 2005. "Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology." *Science* 310: 987-991.

Hagerstrand, T. 1965. "A Monte Carlo Approach to Diffusion." *Archives Europeenes de Sociologie* 6: 43-67.
Hannon, B. 1987. "The Discounting of Concern: A Basis for the Study of Conflict." In *Environmental Economics: The Analysis of a Major Interface*, ed. J. Pillet and T. Murota. Geneva, Switzerland: Leimgruber.
Hannon, B. 1994. "Sense of Place: Geographic Discounting by People, Plants, and Animals." *Ecological Economics* 10: 157-174.
Hannon, B. and M. Ruth. 2001. *Dynamic Modeling*. New York, NY: Springer-Verlag.
Henrickson, L. and B. McKelvey. 2002. "Foundations of "New" Social Science: Institutional Legitimacy from Philosophy, Complexity Science, Postmodernism, and Agent-based Modeling." *Proceedings of the National Academy of Sciences* 99(3): 7288-7295. http://www.pnas.org/cgi/doi/10.1073/pnas.092079799
Holland, J. H. 1998. *Emergence: From Chaos to Order*. Boston, MA: Addison-Wesley Longman Publishing.
Hollingshead, A. B. 1949. *Elmtown's Youth*. New York, NY: Wiley.
Johnson, J. 1981. "Some Structures and Notation of Q-Analysis." *Environment and Planning B* 8: 73-86.
Judy, R. W. and J. M. Lommel. 2002. *21st Century Workforce: Vermilion County*. Indianapolis, IN: The Vermilion County Workforce Investment Board.
Kennedy, J. and R. Eberhart. 2001. *Swarm Intelligence*. San Francisco, CA: Morgan Kaufmann Publishers.
Kenrick, D. T., N. P. Li and J. Butner. 2003. "Dynamical Evolutionary Psychology: Individual Decision Rules and Emergent Social Norms." *Psychological Review* 110(1): 3-26.
Latane, B. 1981. "The Psychology of Social Impact." *American Psychologist* 36: 343-356.
Latane, B. 1996. "Dynamic Social Impact: The Creation of Culture by Communication." *Journal of Communications* 46(4): 13-25.
Laumann, E. O. 1973. *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York, NY: John Wiley and Sons.
Laumann, E. O., P. V. Marsden and J. Galaskiewics. 1977. "Community-Elite Influence Structures." *American Journal of Sociology* 83(3): 594-631.
Laumann, E. O. and F. U. Pappi. 1973. "New Directions in the Study of Community Elites." *American Sociological Review* 38: 212-230.
Lazar, A., D. Chavalarias and T. K. Ahn. 2002. "Endogenous Network Formation and the Evolution of Preferences." In *Agent 2002*. University of Chicago.http://www.econ.iastate.edu/tesfatsi/LazarChavAhn.EndNet.pdf
Ley, D. 1974. *The Black Inner City as Frontier Outpost: Images and Behavior of a Philadelphia Neighborhood*. Washington, DC: The Association of American Geographers.
Macy, M. W., J. A. Kitts and A. Flache. 2003. "Polarization in Dynamic Networks: A Hopfield Model of Emergent Structure." In *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, ed. R. Breiger, K. Carley and P. Pattison. Washington, DC: National Academy Press.
McFadden, D. 1991. "Advances in Computation, Statistical Methods, and Testing of Discrete Choice Models." *Marketing Letters* 2(3): 215-229.
McPherson, M., L. Smith-Lovin and J. M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415-444.

Metcalf, S. S. and M. Paich. 2005. "The Spatial Dynamics of Social Network Evolution." In *Proceedings of the 23rd Interational Conference of the System Dynamics Society*. Boston, MA: System Dynamics Society.http://www.systemdynamics.org/conferences/2005/proceed/papers/METCA487.pdf
Milgram, S. 1967. "The Small World Problem." *Physiology Today* 2: 60-67.
Moore, G. T. and R. G. Golledge, Eds. 1976. *Environmental Knowing*. Stroudsburg, PA: Dowden, Hutchinson and Ross, Inc.
Moore, M. 1989. "Roger and Me." USA: 91 min.
Paez, A. and D. M. Scott. 2007. "Social Influence on Travel Behavior: A Simulation Example of the Decision to Telecommute." *Environment and Planning A* 39(3): 647-665.
Parker, D. C., T. Berger and S. M. Manson, Eds. 2002. *Agent-Based Models of Land-Use and Land-Cover Change: Report and Review of an International Workshop. Irvine, California October 4-7, 2001*. LUCC Report Series No. 6. Bloomington, IN: LUCC International Project Office.
Peake, L. J. 1995. "Toward an Understanding of the Interconnectedness of Women's Lives: the 'Racial' Reproduction of Labor in Low-Income Urban Areas." *Urban Geography* 16(5): 414-439.
Preston, V. and S. M. Taylor. 1981. "Personal Construct Theory and Residential Choice." *Annals of the Association of American Geographers* 71(3): 437-451.
Rainwater, L. 1970. *Behind Ghetto Walls: Black Families in a Federal Slum*. Chicago, IL: Aldine Publishing Co.
Resnick, M. 1994. *Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds*. Cambridge, MA: MIT Press.
Rittel, H. W. J. and M. M. Webber. 1973. "Dilemmas in a General Theory of Planning." *Policy Sciences* 4: 155-169.
Rowe, S. and J. Wolch. 1990. "Social Networks in Time and Space: Homeless Women in Skid Row, Los Angeles." *Annals of the Association of American Geographers* 80(2): 184-204.
Schank, R. C. 1999. *Dynamic Memory Revisited*. Cambridge, UK: Cambridge University Press.
Schelling, T. C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1: 143.
Schelling, T. C. 1978. *Micromotives and Macrobehavior*. New York, NY: W.W. Norton and Company.
Schweitzer, F. 2003. *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences*. New York, NY: Springer.
Shalizi, C. R. 2005. "Methods and Techniques of Complex Systems Science: An Overview." In *Complex Systems Science in Biomedicine*, ed. T. S. Deisboeck and J. Y. Kresh. Dordrecht, The Netherlands: Kluwer Academic Press. http://arxiv.org/PS_cache/nlin/pdf/0307/0307015.pdf
Sterman, J. D. 2000. *Business Dynamics: Systems Thinking and Modeling in a Complex World*. Boston, MA: McGraw-Hill.
Van de Bunt, G. G., M. A. J. Van Duijn and T. A. B. Snijders. 1999. "Friendship Networks Through Time: An Actor-Oriented Dynamic Statistical Network Model." *Computational and Mathematical Organization Theory* 5(2): 167-192.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke and G. Ulfarsson. 2003. "Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim." *Networks and Spatial Economics* 3(1): 43-67.
Wasserman, S. and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.
Watts, D. J. 1999a. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, NJ: Princeton University Press.
Watts, D. J. 1999b. "Networks, Dynamics and the Small World Phenomenon." *American Journal of Sociology* 105(2): 493-527.
Watts, D. J. and S. H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393: 440-442.
Weisbuch, G., G. Deffuant, F. Amblard and J. P. Nadal. 2001. "Interacting Agents and Continuous Opinion Dynamics." *Santa Fe Institute Working Paper 01-11-072*. http://www.santafe.edu/sfi/publications/wpabstract/200111072
Westervelt, J. D. and L. D. Hopkins. 1999. "Modeling Mobile Individuals in Dynamic Landscapes." *International Journal of Geographical Information Science* 13(3): 191-208.
White, D. R. 2004. "Network Analysis and Social Dynamics." *Cybernetics and Systems* 35(2-3): 173-192.
White, D. R., J. Owen-Smith, J. Moody and W. W. Powell. 2004. "Networks, Fields and Organizations: Micro-Dynamics, Scale and Cohesive Embeddings." *Computational and Mathematical Organization Theory* 10(1): 95-117.
Wiegand, T., F. Jeltsch, I. Hanski and V. Grimm. 2003. "Using Pattern-oriented Modeling for Revealing Hidden Information: A Key for Reconciling Ecological Theory and Application." *Oikos* 100(2): 209-415.
Wilson, W. J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago, IL: The University of Chicago Press.
Young, H. P. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.
Young, H. P. 2001. "The Dynamics of Conformity." In *Social Dynamics*, ed. S. N. Durlauf and H. P. Young. Washington, DC: Brookings Institution Press, 133-153.

APPENDICES

Appendix A: Interview Guide

I'm trying to understand what makes a neighborhood. How does a neighborhood form, how does it change? I know that neighborhoods depend on *who* lives there and how they know each other. I'd like to ask you some questions about your own experience with neighborhoods in Danville. I've got a tape recorder to make note-taking easier. If you feel uncomfortable with any of the questions, you can "pass" on them. Your recorded response will be anonymous. Please read over this consent form and sign it. You'll keep a copy and you can contact me at any time.

Are you ready to begin?

---

1. How long have you lived in Danville?
2. Would you mind sharing your age?
3. How long have you lived in this house?
4. Do you rent or own your home?
5. How many people live in your house, including you?
6. Where did you live before?
7. *Why* did you decide to move to this house?
8. Were there any people (friends or family) who influenced your decision? If so, how?
9. Have you ever thought about leaving? If so, when and why?

---

10. How would you describe this neighborhood when you moved here?
11. Did any of your friendships change when you moved here? In what way?
12. If you were to draw your neighborhood boundary at that time, where would it be? Why?
13. Did you know any of your neighbors *before* you came here? *After*?
14. How often did you interact with your neighbors when you moved here?

---

15. *When* did your neighborhood association form?
16. *Why* did your neighborhood association form?
17. *How* did your neighborhood association form? Were you part of that process?

18. How were the boundaries of the association decided?

19. If you could change one thing about the association, what would it be?

20. How has the neighborhood association affected your life here?

21. How would it affect you if the neighborhood association dissolved?

---

22. How many of your neighbors would you consider close friends?

23. How often do you interact with your neighbors? With your friends?

24. In what ways are neighborhood friendships important to you?

25. Who is your closest friend in this neighborhood? In Danville?

26. How would it affect you if your best friend left this neighborhood? If he/she left Danville?

27. How does it affect you when new people move into this neighborhood?

28. Would you consider leaving your neighborhood for another one in Danville? Which?

---

29. Using a reference map of Danville, describe distinct neighborhoods. How are they different from each other?

30. Where do your closest friends and family live?

31. How often do you visit your friends and family? On what occasions?

32. Do you go to church? Which one? Where is it located?

33. How many of your friends and family are part of your church?

34. How would it affect you if the church dissolved?

35. Do you avoid certain places? Which ones? In what way?

I'd like to interview up to ten people for this project. Can you recommend anyone else who might be good to talk to? Someone who lives or used to live in this neighborhood? Or someone from another changing neighborhood in Danville?

Is there anything else you'd like to add? Thank you for your time!

Human Subjects in Research Consent Form
Research on How Neighborhoods Are Made

I will participate in a research study designed to understand how neighborhoods are made from individual choices and social ties to place. This study is conducted by investigator Sara Metcalf at the University of Illinois at Urbana-Champaign. Participants must be 18 years or older.

1. I understand that participants in this study will include individuals who currently or have recently lived in a Danville neighborhood.

2. I understand that the duration of the study will be from summer 2005 through summer 2006, and that most interviews will occur during summer and fall 2005.

3. I understand that I may be asked to respond to interview questions, and my responses may be recorded on audiotape for research purposes. The interviews may take 1 to 1½ hours of my time.

4. I understand that I may refuse to answer any questions that make me feel uncomfortable.

5. I understand that there are no psychological, emotional or physical risks to participation in this study.

6. I understand that I may benefit from this research by gaining an improved understanding of how I influence neighborhoods.

7. I understand that I will not receive any monetary compensation for participation in this study.

8. I understand that I will remain anonymous in any reports of research findings from this study.

9. I understand that participation in this study is voluntary and that I may withdraw at any time.

10. I understand that this research study has been reviewed and approved by the Institutional Review Board—Human Subjects in Research, University of Illinois. For research-related problems or questions regarding subjects' rights, I can contact the Institutional Review Board at 217-333-2670 or email irb@uiuc.edu.

I have read and understand the explanation provided to me. I have had all my questions answered to my satisfaction, and I voluntarily agree to participate in this study.

I have been given a copy of this consent form.

____ ________________________

Subject/Participant Date

____ ________________________

Researcher Date

For more information, contact Sara Metcalf. Phone: 217-390-7421. Email: ssm@uiuc.edu
This research is under the supervision of Dr. Bruce Hannon, Dept. of Geography, University of Illinois.

Appendix B: Neighborhood Research Report-Out

Results of Neighborhood Research
For the Kentucky-Tennessee-Delaware (KTD) Neighborhood Association

Sara S. Metcalf
University of Illinois at Urbana-Champaign
July 13, 2006

For the past year, I've been trying to understand what makes a neighborhood. How does a neighborhood form, how does it change? I know that neighborhoods depend on *who* lives there and how they know each other. I chose Danville as the focus of my research because the overall population is stable, but there have been lots of changes within Danville neighborhoods. John Dreher introduced me to several of the neighborhood associations in town. I first visited the KTD meeting in August 2005, as you were preparing for the block party. The strength of this association impressed me, and after a few more visits I began to interview some of the residents from this area.

For my research, I asked people about their experiences living in different neighborhoods in Danville as well as other communities. I was particularly interested in the influence of social ties (friends and family) on choices of where to live. People choose to move for a lot of different reasons, but having strong connections to people in the place that you live helps to build a solid, stable community.

At this stage, I have finished my interviews and observations of Danville. The city of Danville also provided me with computer data that I've used to look at local migration patterns for homeowners in the past few years. The final stage of my research is to put what I've learned into a computer model that can be used to experiment with different ideas about the role of social ties on neighborhood stability.

What follows is a selection of quotes that relate to the KTD neighborhood experience and the neighborhood association. I have grouped these quotes according to certain themes. I believe that reading your neighbors' words is more powerful than my interpretation of them!

I'm providing this update on what I've learned from the study so far, since I'm moving to Texas to finish the computer part of the project. Thanks so much to everyone who welcomed me here and shared their perspective with me. I hope that you continue to make your neighborhood stronger as you get to know each other through the KTD neighborhood association.

I appreciate any additional comments or concerns that you might have about this study. Please feel free to contact me:

Sara Metcalf
10995 Woodlands Drive
College Station, TX 77845
(979) 764-3925 (home)
(217) 390-7421 (mobile)
sara.metcalf@gmail.com

Concerns and Changes

"Over the years the neighborhood has changed, and we have a younger group coming in now, that's in now, and kids and so forth. When I came in it was so quiet, you know. Now I'm thinking what is all this noise out here, you know. In the summertime the stereos are blasting and all that and cars up and down the street."

"Now the little kids, they are running all over and it scares me to death. I don't say anything but sometimes I think they let them run a little too loose and it's not that safe. It's a main drag."

"They wanted to get the traffic to slow down. To keep kids from running wild in the neighborhood. Maybe to enhance the neighborhood and just in general they got sick and tired of being a deteriorating area of the community and no one really taking an interest."

Past Experiences with Neighborhoods

"I just thought back to my past and how we had our block clubs and our block associations, the camaraderie it brought about and the connections for just neighborhood improvement and well being for our kids."

"I remember as a kid, they used to actually as a neighborhood do things to beautify the neighborhood. It was kind of infectious that when someone was doing something with their home and their property and because of the neighborhood – everybody would do it. So I hope we can get to the point where we can have enough uniformity that we can do things as a neighborhood and spend a little money, our own money on our property and build it up."

KTD Neighbor Experiences

"When I meet a neighbor, I see if there's anything that we have in common. I look to see if there's any need that maybe I could address."

"When my neighbors here went on vacation for about three weeks, I'd get the mail in for them. I've got their cell phone number so I can call them if anything happened. I was never close with them before that – so it is helping to build some bonds here that I think are just wonderful. People are getting to know one another."

"Every time I would move in a neighborhood, if it was somebody I don't know or didn't know before, I would take something like a pie to go welcome."

"I introduced myself because I was working, I was going in and out all the time and there were some situations I wanted them to be aware of. And I wanted to know who lived next door to us and around us."

"When my neighbor first came in, and at the time I had the mentality I'm going to welcome these new folks in, let them know we're here if they need any help with anything. So I went over there and introduced myself. He thanked me, he appreciated it, but he told me that he was a private person; that he was just moving into the neighborhood and his philosophy that he doesn't bother people so he doesn't want people to bother him. I didn't want him to think that we were trying to intrude on his privacy or his new life. Then one day he said, "Do you need any help with anything?" Slowly we developed a kind of quiet friendship. So I would tell him when I'm going to leave, and I'd say if you would kind of keep an eye on the place."

Privacy and Security

"There was an old man that was stabbed. I got leery of the neighbors. I'd speak to those neighbors that I knew, but other people – I didn't reach out to. There's been some things that happened and it makes you leery of who's moving in next door. I believe in God but I believe God wants me to use wisdom and to be careful about things and people."

"What I see with this group here like I said before is a desire to get back to where you could leave your door unlocked. They want that security and watch out and you immediately notice somebody strange."

"A lot of the neighbors just mostly keep to themselves. They do their own thing."

"As we began to see that older age group begin to leave the neighborhood, I kind of took a private type of stance, because as I would see people come moving in – and I didn't know who they were. I didn't know if they were friendly or not. I know a lot of people are so private today."

"If you want to talk to me you can come talk to me. I'm not going to butt into your life."

"Now people have become so private, but I think it's a defense mechanism. You know, they don't want you to know what they have, they don't want you to know too much about them because one they don't want you to come and ask them to do anything you know; and it's unfortunate it's gotten that way but that's the way of life."

Fear

"The association is helping to give me backbone. Fear can just ruin your life. It can paralyze you. I'm not going to live like that."

"I was just getting gas last night and got really frightened. I startle easily. I've had some very off-color remarks made. I don't like to be out at nighttime up in those stores. I just don't feel that safe."

"Somebody tried to rob us. And then every night I would get up just to look out the window. It was like I was scared. It was like we've been invaded."

"They come in with a sense of trying to intimidate, and I think a lot of that comes from their fear of not knowing what to expect, how they're going to be accepted. If I'm going to be accepted badly, I'm going to behave badly."

Prejudice and Bias

"I try to operate on the system that every person is entitled to dignity and respect. Person by person I try to adjust. When you get right down to the nitty gritty everybody has a prejudice and it may not show, but it can be felt."

“I have come to this neighborhood with a lot of prejudice and a lot of biases. I think God is working on humbling me. I’m coming into it with my eyes open. It’s not about what we have. It’s who we are and what we do while we’re here. I think there are lots of reasons why I’m here and in this house and in this community.”

“There may be a few good ones in there and you miss those because of the other stuff, the bad stuff.”

Reporting Trouble

“We’ve got a neighborhood watch going. If we see something going on we can call the police and report it - but at least we can alert one another as to this is going on.”

“I see something going on, I call the police and I don’t worry about it. I’ve always just been afraid to call the police because I’ve had neighbors that have had scanners. Now that they’ve told me, you just tell them you’re with this association, and you tell them you don’t want this over the scanner – and they will do it.”

“I’m not afraid to tell somebody that they are playing their radio too loud. The speed – I’m not afraid to call the cops and tell them to come get them or if there is a lot of loud noises. Whether I confront them directly depends on what I think that person’s personality is. I ache in the morning if I have to wrestle somebody. Sometimes the badge and gun get a little bit more attention than the average Joe.”

Rental Properties

“The association has changed my attitude about our area. You think about your property value and a lot of people don’t own their homes around here. So they aren’t thinking about that stuff, property value.”

“A lot of the homes in this area are rentals, and rentals are going to attract more transient type of people.”

“I would really like to see a better relationship, more activity within the residents because we are so oriented toward rental properties. I think the landowners ought to become more involved, more aware that an effort is being made, and I think that's probably why I became familiar with the landowner next door. Awareness – communication is a real key; and if you don't know who to communicate with, it probably isn't going to happen.”

Challenges

“I don’t know if it was kids or whoever, they spray painted the garage. They wrote it out and then he had wiped it, scrubbed it off. It was like every day, they would put something on the thing. Then I was like, “just leave it alone because they’re going to keep coming back.” It’s still up there. We don’t even mess with it - you don’t know who’s doing these things.”

“My house needs a lot of work on it, but that’s some financial problems that I’ve got.”

"Some of people don't feel a need to go to the meetings. They are satisfied – they take care of their houses, they do this. They are usually older people and others are renters, but they work nights. So they can't."

Awareness

"I encourage the neighbors to just be aware of what is happening."

"It is helping to just get to know the neighbors. It makes a big difference knowing who you live next door to or across the street from."

"I don't think anybody wants anybody in their face, but let them know there is a face."

Building Community

"I consider a neighborhood as where I could come to you, you could come to me, we can help each other out instead of we just going past, hi, and that's it. I don't even know your name."

"To me a neighborhood is something that grows; something that develops; a place where you love to be, you enjoy to be; you bring your family."

"I just want to see the neighbors that are doing this just to stay together and to stay formed and just to have a strong bond there to keep our neighborhood strong."

"We should do more to build relationships with each other, because there tends to be more caring about each other when you have a relationship with each other."

"We have neighbors coming together, leaving contact numbers, getting to know more people in your neighborhood."

"People wouldn't care as much if they didn't know that someone cared. It makes a bond. You need the neighborhood bond to have a better neighborhood."

Visible Changes

"We just want the neighborhood to look good – not like a dump. Things are cleaning up a lot. We've come a long way with it."

"Whether it had to do with the potholes in the streets or an abandoned house being gone or some of the traffic moved out, we didn't care as long as it was something and we could noticeably jot down this is what happened this month and last month this happened. It was really pretty exciting."

"If I show an interest to you in how you living over there and your yard, help you keep it clean, then they'll do it, too. Before I would walk past somebody's house or whatever or the sidewalk and see trash and I just keep walking. But now I do pick up stuff even if it's not on my property."

"My theory is keeping things picked up and respectable and responsible for – has nothing to do with income."

Experiences in the KTD Association

“Before the neighborhood association came about, I never thought about the neighborhood. My neighbors, the person the left and to the right of me, the people directly across the street – probably like this little square area but past that, it's not like we had any really true contact with each other outside of saying "hi" and "goodbye."”

“I felt very much like an outsider. I was needing to be welcomed and embraced. They just had another agenda at the time. I try to look at those things and not take it personally, but that’s how I felt the first time.”

“I used to think the association was like a bickering thing. I can hear that stuff every day all day. I already knew about half of the people they were talking about. But then they started to get the ball rolling. One lady was having a problem, a serious problem that she would be in tears. But then they went down to city hall, talked to the mayor, talked to the police chief, and it’s straightened out. The lady is feeling so much better. I’m glad that happened.”

“I got to meet a few other people that were further down the street that I would talk and connect with better.”

“We've got a connection in that regard on how to bring on more people on the block to be involved so that they can see the benefit in the work that you put into it, especially if you plan to stay in the neighborhood or even to have property improvement and whatever. It's all for the better good. But to have people to come out to understand that it's a benefit for us all, that's the hardest sell, especially to the younger people.”

“I've become more familiar with who the neighbors are, and in that it gives me a little more comfort and knowledge of knowing kind of what might be expected. And I've seen some improvements. People have started to take more pride and connect more with one another.”

“Because of the neighborhood association I’m getting to know them more now. There was a time where we had people move in or move out. They weren’t there long enough to know who they were. And now we have, I think, some neighbors that are fairly stable. But the interaction is not as much. It’s getting better because of the neighborhood association. And I want to know the neighbors, you know. I want the neighbors to know us.”

“I like having the voice of our city officials, because as an association, when you call them they better come. And we’ve got some things done because of that. It gives me an avenue to be involved with the political system with actually not being involved. I like that avenue because I don’t want to have to go down to city hall, I don’t want to have to go to your rallies. We can bring it up at the association. I’ve come to learn that the association, as an association, has a stronger voice than an individual.”

Suggestions for the KTD Association

“I think if we would get together more often, we would know each other more. Then we would trust each other more.”

“At the beginning of each meeting I would have some introductions and maybe even an ice-breaker just to be able to connect on a personal level. Even maybe nametags with our name and address.”

“I would like more involvement. I would like more people coming out and being more hands on actively involved in things. A lot of what we still do is try to bring people out. If we could more equally divide the work, then I think we would grow more and prosper more and not promote burnout. It's a lot of work, but I don't think people truly understand if we divided the work, if everybody had a hands-on involvement and we divided the work, it wouldn't be as much.”

“I think when you come to a neighborhood you need to lay it down. Welcoming packet, these are what you can do and what you can’t do. This is a neighborhood association, would you do be interested in doing this?”

“I would change our process so it’s a little bit more formal. There have been some meetings that we have gone over things two and three times that we already talked about in other meetings. If you have secretary reports and things like that, you can say, you know, we covered that. And you can set yourself action items. I like that because assign somebody to do that and run it down and then it’s done and you can close it out and keep rolling.”

“I would want to get more people involved. Because we have the same kind of group of people. At first there was a lot of people. In the first couple of meetings, when we were talking about getting things done, they wanted them done like the next day. “Why didn’t this get done yesterday?” They don’t realize that, like John Dreher says, everything has to go through a process.”

Overcoming Adversity

“I’m going to do my level best to make it a home and be a part of this community. I’m not going to be chased in and become a prisoner of my home because I encounter some bad behavior on the part of other people. I am going to do my best to seek out people with whom I can develop relationships with and have friendships with, and I think that’s so important.”

“Deterioration of the neighborhood has never been a force to make me want to leave.”

“I think the world has become a place that if you find something wrong with your neighborhood or where you live, the easiest thing to do is leave. The toughest thing to do is to stay there and make it better. I read a little quote that says, “adversity does not build character, it reveals it.” I like that. And what I hope is over time, if we’ve had adversity in this community, it’s not going to build character for the community but it’s going to reveal the character that we already have. That’s my hope for this association because we have good people, we have strong leaders, and they’re not made; it’s already there; it’s just being brought out.”

Appendix C: Metadata for Parcel Dataset

Horizontal coordinate system

Projected coordinate system name: NAD_1927_StatePlane_Illinois_East_FIPS_1201

Geographic coordinate system name: GCS_North_American_1927

Details

Grid Coordinate System Name: State Plane Coordinate System 1927

SPCS Zone Identifier: 1201

Transverse Mercator Projection

Scale Factor at Central Meridian: 0.999975
Longitude of Central Meridian: -88.333333
Latitude of Projection Origin: 36.666667
False Easting: 500000.000000
False Northing: 0.000000

Planar Coordinate Information

Planar Distance Units: survey feet

Coordinate Encoding Method: coordinate pair

Coordinate Representation

Abscissa Resolution: 0.000128

Ordinate Resolution: 0.000128

Geodetic Model

Horizontal Datum Name: North American Datum of 1927

Ellipsoid Name: Clarke 1866

Semi-major Axis: 6378206.400000

Denominator of Flattening Ratio: 294.978698

Bounding coordinates

Horizontal

In decimal degrees

West: -87.679502
East: -87.528728
North: 40.227759
South: 40.086123

In projected or local coordinates

Left: 682947.312495
Right: 724674.687486
Top: 1297527.125007
Bottom: 1246270.000015

Appendix D: Example Raw (Unstructured) Data Files for Model Objects

Table 9. Row Excerpt from Parcel Data (13166parcels.csv)

HouseID	X	Y	BlockName	BlockGroupName	PARCEL_ID
0	696788.25	1262927.79	171830001001021	171830001001	2969
1	695772.30	1259700.80	171830001002017	171830001002	1394
2	695846.21	1259800.77	171830001002017	171830001002	1395
3	699342.50	1257996.72	171830001002039	171830001002	801
4	699354.35	1258067.47	171830001002039	171830001002	813
5	702442.85	1256717.10	171830003002014	171830003002	3552
6	707106.23	1269468.04	171830004001010	171830004001	4237
7	708410.46	1270161.17	171830004001010	171830004001	4311
8	707935.89	1269434.92	171830004001010	171830004001	4316
9	706175.22	1269669.20	171830004001010	171830004001	4736
10	709026.43	1268988.14	171830004001010	171830004001	4798
11	705632.51	1268624.17	171830004001010	171830004001	4936
12	705417.17	1268617.81	171830004001010	171830004001	4939
13	705040.87	1268609.17	171830004001010	171830004001	4949
14	704617.36	1268594.16	171830004001010	171830004001	4956
15	706186.89	1268805.73	171830004001010	171830004001	5011
16	704883.19	1269196.26	171830004001010	171830004001	5028
17	704008.35	1268344.67	171830004001010	171830004001	5282
18	704274.89	1268277.71	171830004001010	171830004001	5480
19	704431.87	1268336.88	171830004001010	171830004001	5523
20	706863.05	1268243.72	171830004001010	171830004001	5563
21	704433.25	1268196.73	171830004001010	171830004001	5571
22	706094.13	1268062.77	171830004001010	171830004001	5679
23	705959.20	1268058.08	171830004001010	171830004001	5684
24	710047.18	1268516.30	171830004001014	171830004001	5352
25	709083.23	1266953.29	171830004001022	171830004001	700
26	709018.58	1266944.83	171830004001022	171830004001	707
27	708956.80	1266942.42	171830004001022	171830004001	711
28	708894.44	1266939.07	171830004001022	171830004001	712
29	708832.41	1266936.68	171830004001022	171830004001	716
30	708771.48	1266933.99	171830004001022	171830004001	720
31	708710.30	1266932.07	171830004001022	171830004001	723
32	708647.00	1266930.44	171830004001022	171830004001	727
33	708589.61	1266928.71	171830004001022	171830004001	730
34	708516.92	1266944.40	171830004001022	171830004001	733
35	708580.82	1267454.21	171830004001022	171830004001	3582
36	709552.40	1268121.71	171830004001022	171830004001	5592
37	709343.40	1268114.57	171830004001022	171830004001	5600
38	709001.85	1267996.86	171830004001022	171830004001	5608
39	708575.04	1267969.71	171830004001022	171830004001	5748
40	708783.98	1267404.67	171830004001022	171830004001	5804
41	709542.05	1267932.54	171830004001022	171830004001	5806
42	708492.80	1267890.48	171830004001022	171830004001	5876
43	709569.58	1267840.88	171830004001022	171830004001	5891
44	708602.82	1267771.87	171830004001022	171830004001	5960
45	708919.55	1267692.58	171830004001022	171830004001	5990

Table 10. Row Excerpt from Owner Data File (7576owners.csv)

OwnID	HouseID	MoveCode	FAMNAME	FamCount
8	47	Left	UPPERMAN	1
9	48	Stayed	LANCON	1
10	52	Stayed	JACKSON	14
11	56	Stayed	STEWART	9
12	57	Stayed	PARKER	14
13	59	Stayed	ROBINSON	20
14	60	Stayed	POPE	5
15	62	Stayed	TAPIA	2
16	64	Stayed	BUTLER	15
17	65	Stayed	GARZA	4
18	68	Stayed	WEST	7
19	70	Stayed	ELLIS	17
20	72	Stayed	PEREZ	18
21	73	Stayed	COX	19
22	77	Left	KIDWELL	5
23	78	Left	SANKS	2
24	79	Stayed	SCHNEIDER	3
25	80	Stayed	COLLIER	4
26	81	Left	WILLIAMS	44
27	84	Stayed	MORGAN	13
28	87	Stayed	LORENZ	1
29	88	Left	SMITH	99
30	89	Stayed	DAVIS	73
31	91	Stayed	BRUNS	1
32	92	Stayed	CITIZEN	2
33	93	Left	BAKER	8
34	94	Stayed	SHADLEY	1
35	95	MovedFrom	WILLIAMSO	9
36	96	Stayed	BLEDSOE	1
37	97	Stayed	GRIFFITH	3
38	99	Stayed	HARRINGTO	1
39	100	Stayed	LUCAS	11
40	101	Stayed	GEADES	2
41	103	Left	KNIGHT	7
42	108	Stayed	KEGLEY	3
43	109	Left	GARMAN	2
44	112	Stayed	HANER	2
45	113	Stayed	BAKER	8
46	114	Stayed	BRADFORD	1
47	116	Stayed	BAYS	2
48	117	Stayed	CROOM	1
49	118	Stayed	CROOK	3
50	119	Stayed	ESTRADA	1
51	120	Stayed	WILSON	32
52	122	Left	RODRIQUEZ	1
53	132	Stayed	PAXTON	3
54	133	Stayed	RAMIREZ	3
55	135	Stayed	WILLIAMS	44

Table 11. Row Excerpt from Block Data File (764blocks.csv)

BlockID	BlockName	BlockGroupName	Parcels	Owners	Renters	Vacant	OwnWprob	RentWprob	CentroidX	CentroidY
0	171830001001000	171830001001	11	0	5	6	1.00	0.50	697203.11	1263745.78
1	171830001001003	171830001001	8	1	2	5	1.00	0.00	696466.22	1263725.16
2	171830001001004	171830001001	3	0	2	1	0.00	0.13	696135.78	1263678.68
3	171830001001005	171830001001	13	0	10	3	0.00	0.38	695705.50	1263298.51
4	171830001001007	171830001001	28	3	19	6	0.83	0.33	695299.28	1262523.27
5	171830001001008	171830001001	32	9	13	10	0.50	0.45	695705.58	1262549.73
6	171830001001009	171830001001	16	0	16	0	0.00	0.00	696098.46	1262568.80
7	171830001001010	171830001001	18	0	15	3	0.00	0.55	696089.15	1263333.61
8	171830001001012	171830001001	5	0	5	0	0.00	0.78	696762.37	1263330.01
9	171830001001013	171830001001	4	0	4	0	1.00	0.64	696943.71	1263331.44
10	171830001001014	171830001001	10	0	10	0	1.00	0.70	697212.94	1263333.15
11	171830001001015	171830001001	7	4	0	3	0.00	0.00	697212.63	1262914.58
12	171830001001016	171830001001	5	0	3	2	0.00	0.00	697478.24	1262916.69
13	171830001001018	171830001001	5	2	3	0	0.00	0.00	697690.40	1262663.36
14	171830001001019	171830001001	4	0	2	2	0.00	0.20	697480.40	1262643.62
15	171830001001020	171830001001	15	9	5	1	0.00	0.74	697215.44	1262395.77
16	171830001001021	171830001001	14	1	12	1	1.00	0.60	696880.61	1262574.99
17	171830001001024	171830001001	8	0	5	3	1.00	0.82	696101.51	1261785.57
18	171830001001025	171830001001	14	3	7	4	0.80	0.50	695729.92	1261767.29
19	171830001001026	171830001001	9	0	5	4	0.00	0.95	695459.49	1261780.73
20	171830001001028	171830001001	4	1	3	0	1.00	1.00	695271.74	1261368.92
21	171830001001029	171830001001	8	2	3	3	0.50	0.93	695520.33	1261358.07
22	171830001001030	171830001001	10	0	10	0	0.00	0.75	695786.01	1261337.09
23	171830001001031	171830001001	7	0	5	2	0.00	0.75	696094.61	1261381.98
24	171830001001033	171830001001	2	0	1	1	0.00	0.76	696893.61	1261426.85
25	171830001001036	171830001001	17	2	12	3	0.00	0.08	697585.39	1261792.64
26	171830001001037	171830001001	11	4	0	7	0.00	0.00	697692.02	1262308.41
27	171830001001038	171830001001	9	4	0	5	0.00	0.00	697484.59	1262288.91
28	171830001001039	171830001001	3	1	1	1	0.00	0.00	697592.28	1261374.04
29	171830001001044	171830001001	11	1	1	9	1.00	0.67	696115.28	1260894.41
30	171830001001045	171830001001	10	0	8	2	1.00	0.57	695789.88	1260845.16
31	171830001001046	171830001001	18	4	10	4	1.00	0.80	695524.11	1260861.13
32	171830001001047	171830001001	16	1	12	3	1.00	0.73	695257.17	1260867.98
33	171830001002000	171830001002	18	7	5	6	0.71	0.83	699684.19	1259605.85
34	171830001002001	171830001002	21	11	6	4	0.44	0.42	699299.78	1259571.11
35	171830001002002	171830001002	8	0	8	0	0.50	0.50	698890.38	1259719.58
36	171830001002003	171830001002	11	0	5	6	1.00	0.33	698495.26	1259771.45
37	171830001002004	171830001002	14	2	12	0	0.00	1.00	698127.92	1259579.90
38	171830001002009	171830001002	15	1	14	0	1.00	1.00	697594.58	1260872.08
39	171830001002015	171830001002	16	1	4	11	0.00	0.67	695654.22	1260177.69
40	171830001002017	171830001002	6	0	6	0	0.00	0.80	695686.56	1259541.66
41	171830001002024	171830001002	20	5	11	4	1.00	0.77	698428.01	1259062.40
42	171830001002025	171830001002	15	5	10	0	0.86	0.83	698508.96	1259441.94
43	171830001002026	171830001002	4	1	3	0	0.00	0.90	698939.22	1259373.16
44	171830001002027	171830001002	7	1	0	6	1.00	0.00	698931.56	1259076.07
45	171830001002028	171830001002	24	9	13	2	1.00	0.45	699511.41	1259071.80
46	171830001002029	171830001002	5	4	0	1	1.00	1.00	699518.87	1258742.85
47	171830001002031	171830001002	13	6	4	3	0.86	0.00	698447.94	1258763.59
48	171830001002036	171830001002	12	5	4	3	0.67	0.50	698438.18	1258430.46
49	171830001002037	171830001002	20	11	7	2	0.91	0.86	698960.70	1258428.52
50	171830001002038	171830001002	24	18	4	2	1.00	0.80	699523.24	1258406.57
51	171830001002039	171830001002	24	13	7	4	0.80	0.57	699538.82	1258013.80
52	171830001002040	171830001002	19	13	3	3	0.92	0.45	698558.70	1257636.34
53	171830001002041	171830001002	18	5	11	2	1.00	0.93	698451.69	1258028.79
54	171830001002045	171830001002	15	13	1	1	0.91	1.00	699573.68	1257682.43
55	171830001002046	171830001002	27	20	7	0	0.88	1.00	699961.22	1256855.02
56	171830002001000	171830002001	8	3	0	5	0.00	0.00	701594.83	1263969.79
57	171830002001001	171830002001	8	5	2	1	0.00	0.00	701277.10	1263958.36
58	171830002001002	171830002001	9	1	2	6	0.00	0.00	700920.59	1263937.06
59	171830002001005	171830002001	6	4	2	0	0.33	1.00	699966.99	1263493.08
60	171830002001006	171830002001	15	4	0	11	0.20	0.00	700270.39	1263522.01
61	171830002001007	171830002001	14	3	6	5	0.33	0.25	700589.62	1263533.93
62	171830002001008	171830002001	16	10	2	4	0.00	0.50	700929.60	1263518.41
63	171830002001009	171830002001	16	7	7	2	0.14	0.40	701277.31	1263530.15
64	171830002001010	171830002001	15	5	7	3	0.00	0.00	701598.38	1263532.75
65	171830002001011	171830002001	15	6	4	5	0.14	0.50	701607.51	1263017.38

Table 12. Column Excerpt[20] from BlockGroup Data File (28blockGroups.csv)

BlockGroupID	BlockGroupName	ChildProb	WhiteChild	NonWhiteChild	Stayed	MovedNear	MovedFar	Inc1	Inc16	White1	White16	Nonwhite1	Nonwhite16
0	171830001001	0.1237	0.1205	0.1290	0.2643	0.5495	0.1863	0.3651	0.0000	0.4104	0.0000	0.2903	0.0000
1	171830001002	0.2453	0.2295	0.2949	0.4231	0.4458	0.1311	0.1646	0.0000	0.1189	0.0000	0.3077	0.0000
2	171830002001	0.2677	0.0000	0.3063	0.6738	0.1943	0.1319	0.2283	0.0000	0.2188	0.0000	0.2297	0.0000
3	171830002002	0.1633	0.0000	0.1633	0.6897	0.2808	0.0296	0.4694	0.0000	0.4694	0.0000	0.4694	0.0000
4	171830002003	0.3590	0.3071	0.5634	0.5000	0.3491	0.1509	0.1054	0.0171	0.1107	0.0000	0.0845	0.0845
5	171830003001	0.3051	0.2623	0.3695	0.6480	0.3079	0.0441	0.1417	0.0000	0.0951	0.0000	0.2118	0.0000
6	171830003002	0.3924	0.4129	0.3176	0.6277	0.3203	0.0519	0.0785	0.0000	0.0871	0.0000	0.0471	0.0000
7	171830004001	0.5656	0.4344	0.6038	0.3313	0.4847	0.1840	0.4861	0.0074	0.2459	0.0000	0.5561	0.0095
8	171830004002	0.3306	0.3109	0.3871	0.5616	0.2785	0.1600	0.0569	0.0153	0.0412	0.0206	0.1022	0.0000
9	171830005001	0.3230	0.2745	0.5269	0.5173	0.3017	0.1810	0.1508	0.0000	0.0961	0.0000	0.3808	0.0000
10	171830005002	0.1880	0.0809	0.4918	0.6535	0.2977	0.0488	0.0470	0.0385	0.0636	0.0520	0.0000	0.0000
11	171830006001	0.3984	0.3538	0.4483	0.4960	0.3291	0.1749	0.1138	0.0000	0.0538	0.0000	0.1810	0.0000
12	171830006002	0.1900	0.1656	0.2464	0.6566	0.2956	0.0477	0.2969	0.0000	0.3313	0.0000	0.2174	0.0000
13	171830006003	0.3649	0.3120	0.4142	0.4436	0.4214	0.1350	0.2452	0.0000	0.1600	0.0000	0.3246	0.0000
14	171830007001	0.2862	0.2500	0.8250	0.6437	0.2302	0.1261	0.0283	0.0629	0.0302	0.0336	0.0000	0.5000
15	171830007002	0.2396	0.2207	1.0000	0.5968	0.2532	0.1500	0.0945	0.0242	0.0968	0.0248	0.0000	0.0000
16	171830007003	0.2632	0.2545	0.3939	0.5492	0.3142	0.1366	0.0207	0.0000	0.0220	0.0000	0.0000	0.0000
17	171830008001	0.3661	0.3300	0.6410	0.4013	0.3571	0.2416	0.1518	0.0000	0.1717	0.0000	0.0000	0.0000
18	171830008002	0.2529	0.2584	0.1633	0.6018	0.2958	0.1024	0.0765	0.0059	0.0699	0.0062	0.1837	0.0000
19	171830009001	0.2429	0.2469	0.2090	0.5608	0.3419	0.0973	0.2429	0.0063	0.2078	0.0071	0.5373	0.0000
20	171830009002	0.2901	0.2764	0.5000	0.6145	0.2345	0.1511	0.1225	0.0000	0.1152	0.0000	0.2353	0.0000
21	171830010002	0.3680	0.3339	0.5068	0.6253	0.3318	0.0430	0.1280	0.0093	0.1279	0.0116	0.1284	0.0000
22	171830010003	0.2745	0.2934	0.2321	0.4789	0.4327	0.0884	0.1338	0.0000	0.0998	0.0000	0.2098	0.0000
23	171830011001	0.4176	0.4098	0.4333	0.4502	0.3321	0.2177	0.0769	0.0000	0.1148	0.0000	0.0000	0.0000
24	171830012001	0.1403	0.1345	0.2083	0.4771	0.3866	0.1363	0.2529	0.0000	0.2619	0.0000	0.1458	0.0000
25	171830012002	0.2524	0.2491	0.3111	0.5895	0.2587	0.1518	0.0443	0.0000	0.0468	0.0000	0.0000	0.0000
26	171830013001	0.2481	0.2457	0.2714	0.5537	0.2948	0.1514	0.0363	0.0233	0.0335	0.0207	0.0643	0.0500
27	171830105001	0.2010	0.2010	0.0000	0.6955	0.2142	0.0902	0.1154	0.0000	0.1154	0.0000	0.1154	0.0000

[20] Only the 1st and 16th of the 16 income brackets are shown – for the entire BlockGroup, for the White racial category, and for the NonWhite racial category.

Appendix E: Observed Migration Patterns

Table 13. Owner Migration by BlockGroup[21] from 2001 to 2003

index	FromBG	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
0	171830012001										1																			12
1	171830007001		1								1	1					5								1	1				86
2	171830002001			1																										32
3	171830006001					1																								20
4	171830012002		1			2					4						6		1							1	1	1		117
5	171830003001					1					2					1	1								1		1		1	52
6	171830006002		1																											26
7	171830002002																													11
8	171830008002																	1												20
9	171830005001	1	4			5					6						8													140
10	171830010003					1					1		1								1	1				1				90
11	171830009001												1				1							1				1		50
12	171830008001		3					1			1																			38
13	171830005002						1								1			1												23
14	171830002003		1								1		1												1					41
15	171830013001		4														7													133
16	171830007003		3							1	1	1		1			4	1												81
17	171830007002		2			1					1						3		1											66
18	171830011001																													7
19	171830003002		1									1																		53
20	171830105001																													7
21	171830001001																													10
22	171830009002					1													1											35
23	171830004002		1			1				1				1		1			1						1			1		87
24	171830001002					1												1												36
25	171830006003																	1	1											51
26	171830010002					2								1					1						1					93
27	171830004001																	1												12
28	OUTSIDE	23	79	20	22	134	58	28	2	19	150	91	46	44	21	45	123	94	79	7	62	5	15	52	80	36	49	79	9	

[21] BlockGroups are ordered according to the HashMap index at left. This index is generated in the process of HashMap formation and is used as a key to link to each BlockGroup model object. This re-ordering (in contrast to the sorting by name in Table 12) is important for model initialization (see Main Active Object Class Code in Appendix G).

Table 14. Owner Migration by BlockGroup from 2003 to 2005

index	FromBG	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
0	171830012001																													6
1	171830007001		2			1											2													54
2	171830002001						1				1						1													33
3	171830006001						1																							15
4	171830012002		2			3											5		1											97
5	171830003001						2						1				1									1				51
6	171830006002						1	2											1											30
7	171830002002																													10
8	171830008002													1																12
9	171830005001		1			1		1			2						5	2												110
10	171830010003	1				1		1			1						1		1	1					1	2				58
11	171830009001		1								1																			33
12	171830008001										2																1			27
13	171830005002											1			3															20
14	171830002003		1			1					1																			30
15	171830013001		2			1				1		3					6													92
16	171830007003			1		3					3						3	1												60
17	171830007002		2			1					1																			38
18	171830011001																				1									12
19	171830003002					1					1																			38
20	171830105001																													3
21	171830001001																								1					10
22	171830009002																													28
23	171830004002		2			2		1									2	1							1					56
24	171830001002																2				1									21
25	171830006003															1														37
26	171830010002		2			2												1										2		69
27	171830004001										1			1				1												13
28	OUTSIDE	19	51	20	9	98	44	20	9	15	129	59	23	22	13	25	117	64	45	6	37	4	10	25	44	17	39	70	7	

Appendix F: Java Classes for Data Processing

Processing Package

CreateDataStructure.java

```
package processing;

import io.Serializer;
import io.TextFile;

import java.io.IOException;
import java.util.HashMap;

import structure.BlockContainer;
import structure.BlockGroupContainer;
import structure.CensusMoves;
import structure.DataContainer;
import structure.OwnerContainer;
import structure.ParcelContainer;
import structure.ParcelStatus;

public class CreateDataStructure {
    /*
    * @param args: 28blockgroups.csv, 764blocks.csv, 13166parcels.csv, 7576owners.csv,
    * dataStructure.ser
    * @throws IOException
    */
    public static void main(String[] args) throws IOException {
        //create the Data Container and add the hashmap
        DataContainer dc = new DataContainer();
        HashMap<String,BlockGroupContainer> blockGroups = new HashMap<String,BlockGroupContainer>
        ();
        HashMap<String,BlockContainer> blocks = new HashMap<String,BlockContainer>();
        HashMap<String,ParcelContainer> parcels = new HashMap<String,ParcelContainer>();
        HashMap<String,OwnerContainer> owners = new HashMap<String,OwnerContainer>();
        dc.blockGroups = blockGroups;
        dc.blocks = blocks;
        dc.parcels = parcels;
        dc.owners = owners;
        //get the blockGroups.csv Data file and throw away the header row
        TextFile in1 = new TextFile( args[0] );
        in1.remove(0);
        for(int i=0; i<in1.size(); i++){
            String[] row = in1.getRowArray(i);
            BlockGroupContainer bgc = new BlockGroupContainer();
            //field 0 is id, field 1 is "BlockGroupName" 12-char string
            bgc.key = row[1].trim();
            //create Hashmap
            blockGroups.put(bgc.key, bgc);
            //fields 2-4 are child fractions among total, white and nonwhite HHs
            bgc.tcf = Double.parseDouble(row[2].trim());
            bgc.wcf = Double.parseDouble(row[3].trim());
            bgc.ncf = Double.parseDouble(row[4].trim());
            //now we're on field 5, "Stayed"
            bgc.censusMoves = new CensusMoves(Double.parseDouble(row[5].trim()),
                    Double.parseDouble(row[6].trim()),
                    Double.parseDouble(row[7].trim()));
            //now we're on field 8
            for(int j=0; j<16; j++){
                bgc.incomeGroup[j]=Double.parseDouble(row[j+8].trim());
                bgc.whiteIncome[j]=Double.parseDouble(row[j+24].trim());
                bgc.nonwhiteIncome[j]=Double.parseDouble(row[j+40].trim());
            }
        }//blockGroup loop
        //import blocks.csv Data File as args[1]
        TextFile in2 = new TextFile(args[1]);
        in2.remove(0);
        for(int i=0; i<in2.size(); i++){
            String[] row = in2.getRowArray(i);
            BlockContainer bc = new BlockContainer();
```

```
            bc.key = row[1].trim(); //blockName
            blocks.put(bc.key, bc); //create hashMap
            bc.bgName = row[2].trim(); //blockGroupName, aka bgc.key
            bc.parcelNum = Integer.parseInt(row[3].trim());
            bc.parcelStatus = new ParcelStatus(Integer.parseInt(row[4].trim()),
                    Integer.parseInt(row[5].trim()),
                    Integer.parseInt(row[6].trim()));
            bc.owp = Double.parseDouble(row[7].trim());
            bc.rwp = Double.parseDouble(row[8].trim());
            bc.cenX = Double.parseDouble(row[9].trim());
            bc.cenY = Double.parseDouble(row[10].trim());
        }
        //import parcels.csv Data File as args[2]
        TextFile in3 = new TextFile(args[2]);
        in3.remove(0);
        for(int i=0; i<in3.size(); i++){
            String[] row = in3.getRowArray(i);
            ParcelContainer pc = new ParcelContainer();
            pc.key = row[0].trim(); //parcelID, same as i
            parcels.put(pc.key, pc); //create hashMap
            pc.xPos = Double.parseDouble(row[1].trim());
            pc.yPos = Double.parseDouble(row[2].trim());
            pc.bName = row[3].trim(); //blockName, aka bc.key
            pc.bgName = row[4].trim(); //blockGroupName, aka bgc.key
            pc.pCode = row[5].trim(); //parcel_ID for GIS reference
        }
        TextFile in4 = new TextFile(args[3]);
        in4.remove(0);
        for(int i=0; i<in4.size(); i++){
            String[] row = in4.getRowArray(i);
            OwnerContainer oc = new OwnerContainer();
            oc.key = row[0].trim(); //ownerID, same as i
            owners.put(oc.key, oc); //create hashMap
            oc.pName = row[1].trim(); //matches pc.key
            oc.moveCode = row[2].trim();
            oc.lastName = row[3].trim();
            oc.lastNameCt = Integer.parseInt(row[4].trim());
        }
        //assign lists
        for (String key: dc.parcels.keySet()){
            ParcelContainer pc = dc.parcels.get(key);
            if(dc.blocks.containsKey(pc.bName)){
                BlockContainer bc = dc.blocks.get(pc.bName);
                bc.parcelList.add(pc);
            }
        }//end parcel loop
        for(String key: dc.owners.keySet()){
            OwnerContainer oc = dc.owners.get(key);
            if(dc.parcels.containsKey(oc.pName)){
                ParcelContainer pc = dc.parcels.get(oc.pName);
                pc.owner = oc;
            }
        }//end owner loop
        for(String key: dc.blocks.keySet()){
            BlockContainer bc = dc.blocks.get(key);
            BlockGroupContainer bgc = dc.blockGroups.get(bc.bgName);
            bgc.blockList.add(bc);
        }//end block loop

        //args[4] is dataStructure.ser as a persistent data structure
        Serializer.store(dc, args[4]);
    }
}
```

CreateModelInstances.java

```
package processing;

import io.Serializer;
```

```
import java.io.FileNotFoundException;
import java.io.IOException;
import java.util.ArrayList;

import structure.BlockContainer;
import structure.BlockGroupContainer;
import structure.DataContainer;
import structure.OwnerContainer;
import structure.ParcelContainer;

import caseTown.Block;
import caseTown.BlockGroup;
import caseTown.Household;
import caseTown.ModelObjects;
import caseTown.Parcel;

public class CreateModelInstances {

    /**
     * @param args: dataStructure.ser, modelObjectsFinal.ser
     * @throws IOException
     * @throws FileNotFoundException
     * @throws ClassNotFoundException
     */
    public static void main(String[] args) throws FileNotFoundException, IOException,
    ClassNotFoundException {
        // TODO Auto-generated method stub
        ArrayList<BlockGroup> blockGroupList = new ArrayList<BlockGroup>();
        ArrayList<Block> blockList = new ArrayList<Block>();
        ArrayList<Parcel> parcelList = new ArrayList<Parcel>();
        ArrayList<Household> householdList = new ArrayList<Household>();
        ArrayList<Household> ownerList = new ArrayList<Household>();
        ArrayList<Household> renterList = new ArrayList<Household>();
        ArrayList<Parcel> vacantParcels = new ArrayList<Parcel>();

        DataContainer dc = (DataContainer) Serializer.load(args[0]);

        for (String bgKey: dc.blockGroups.keySet()){
            BlockGroupContainer bgc = dc.blockGroups.get(bgKey);
            BlockGroup bg = new BlockGroup();
            blockGroupList.add(bg);
            bg.name = bgKey;
            bg.incomeGroup = bgc.incomeGroup;
            bg.whiteIncome = bgc.whiteIncome;
            bg.nonwhiteIncome = bgc.nonwhiteIncome;
            bg.wcf = bgc.wcf;
            bg.ncf = bgc.ncf;
            for (BlockContainer bc: bgc.blockList){
                Block b = new Block();
                b.name = bc.key;
                b.xCen = bc.cenX;
                b.yCen = bc.cenY;
                b.parcelStatus = bc.parcelStatus;
                int r = b.parcelStatus.rent;
                int v = b.parcelStatus.vacant;
                b.myBlockGroup = bg;
                blockList.add(b);
                bg.blockList.add(b);
                for (ParcelContainer pc: bc.parcelList){
                    Parcel p = new Parcel();
                    p.name = pc.key;
                    p.myBlock = b;
                    p.xPos = pc.xPos;
                    p.yPos = pc.yPos;
                    b.xAvg += p.xPos;
                    b.yAvg += p.yPos;
                    parcelList.add(p);
                    b.parcelList.add(p);
                    OwnerContainer oc = pc.owner;
                    double rand = Math.random();
                    if (oc != null){
```

```
                        Household h = new Household();
                        h.name = oc.key;
                        h.myParcel = p;
                        p.myHousehold = h;
                        p.ownerOccupied = true;
                        if (bc.owp > rand) h.white = true;
                        householdList.add(h);
                        ownerList.add(h);
                        bg.householdList.add(h);
                    }
                    else if (r > 0) {
                        if (r > v){
                            Household h = new Household();
                            h.myParcel = p;
                            p.myHousehold = h;
                            if (bc.rwp > rand) h.white = true;
                            householdList.add(h);
                            renterList.add(h);
                            bg.householdList.add(h);
                            r = r-1;
                        }//end if rentals dominate
                        else v = v-1; //parcel stays empty
                    }//end if rentals exist
                    if (p.myHousehold == null) vacantParcels.add(p);
                }//end parcel loop
                b.xAvg = b.xAvg/b.parcelList.size();
                b.yAvg = b.yAvg/b.parcelList.size();
            }//end block loop
        }//end blockGroup loop

        ModelObjects mo = new ModelObjects();
        mo.blockGroupList = blockGroupList;
        mo.blockList = blockList;
        mo.parcelList = parcelList;
        mo.householdList = householdList;
        Serializer.store(mo, args[1]); //output serializer filename
    }
}
```

BlockDistance.java

```
package processing;

import io.Serializer;

import java.io.FileNotFoundException;
import java.io.IOException;
import java.io.Serializable;

import caseTown.Block;
import caseTown.ModelObjects;

public class BlockDistance implements Serializable {
    private static final long serialVersionUID = 1L;

    public static void addDistances (ModelObjects mo){
        int size = mo.blockList.size();
        mo.blockDistance = new double[size][];
        for(int row = 0; row < size; row++){
            Block fromB = (Block) mo.blockList.get(row);
            fromB.seq = row;
            double fromX = fromB.xAvg;
            double fromY = fromB.yAvg;
            mo.blockDistance[row] = new double[row+1];
            for(int col = 0; col < row+1; col++){
                Block toB = (Block) mo.blockList.get(col);
                double toX = toB.xAvg;
                double toY = toB.yAvg;
                double xDist = fromX - toX;
                double yDist = fromY - toY;
```

```
                double d = Math.hypot(xDist, yDist);
                mo.blockDistance[row][col]=d;
            }
        }
    }
    /**
     * @param args: modelObjectsFinal.ser
     * @throws ClassNotFoundException
     * @throws IOException
     * @throws FileNotFoundException
     */
    public static void main(String[] args) throws FileNotFoundException, IOException,
    ClassNotFoundException {
        // TODO Auto-generated method stub
        ModelObjects mo = (ModelObjects)Serializer.load(args[0]);
        addDistances(mo);
        Serializer.store(mo,args[0]);
    }
}
```

AssignChildrenIncome.java

```
package processing;

import io.Serializer;

import java.io.FileNotFoundException;
import java.io.IOException;
import java.io.Serializable;

import caseTown.BlockGroup;
import caseTown.Household;
import caseTown.ModelObjects;

public class AssignChildrenIncome implements Serializable {

    private static final long serialVersionUID = 1L;

    public static void addAttributes (ModelObjects mo){
        int bgs = mo.blockGroupList.size();
        double[] tip = new double[16];
        double[] wip = new double[16];
        double[] nip = new double[16];
        double[] max = {10,15,20,25,30,35,40,45,50,60,75,100,125,150,200,250};
        double[] min = {5,10,15,20,25,30,35,40,45,50,60,75,100,125,150,200};
        for (int index=0; index<bgs; index++){
            BlockGroup bg = (BlockGroup) mo.blockGroupList.get(index);
            double[] tir = bg.incomeGroup;
            double[] wir = bg.whiteIncome;
            double[] nir = bg.nonwhiteIncome;
            for (int j=0; j<16; j++){
                for (int k=0; k<j+1; k++){
                    tip[j] += tir[k];
                    wip[j] += wir[k];
                    nip[j] += nir[k];
                }
            }//assign cumulative probs for each income group
            for (int j=0; j<16; j++){
                tip[j] = tip[j]*1/tip[15];
                wip[j] = wip[j]*1/wip[15];
                nip[j] = nip[j]*1/nip[15];
            }//scale so they wind up at exactly one
            int size = bg.householdList.size();
            for (int i=0; i<size; i++){
                //int randomIndex = (int) Math.random()*(size-1);
                Household h = (Household) bg.householdList.get(i);
                double r = Math.random();
                if (h.white && (bg.wcf>r)) h.children = true;
                if (!h.white && (bg.ncf>r)) h.children = true;
                double minP = 0;
```

```
            double maxP = 1;
            boolean[] fit = new boolean[16];
            for (int k=0; k<16; k++){
                if (h.white) maxP = wip[k]; else maxP = nip[k];
                if ((r<maxP)&&(r>=minP)){
                    fit[k]=true;
                    h.income = Math.random()*(max[k]-min[k])+min[k];
                }
                if (h.white) minP = wip[k]; else minP = nip[k];
            }//find the fitting category
        }
    }
}
/**
 * @param args
 * @throws ClassNotFoundException
 * @throws IOException
 * @throws FileNotFoundException
 */
public static void main(String[] args) throws FileNotFoundException, IOException,
ClassNotFoundException {
    // TODO Auto-generated method stub
    ModelObjects mo = (ModelObjects)Serializer.load(args[0]);
    addAttributes(mo);
    Serializer.store(mo,args[0]);
}
}
```

Structure Package

DataContainer.java

```
package structure;

import java.io.Serializable;
import java.util.HashMap;

public class DataContainer implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public HashMap<String, BlockGroupContainer> blockGroups;
    public HashMap<String, BlockContainer> blocks;
    public HashMap<String, ParcelContainer> parcels;
    public HashMap<String, OwnerContainer> owners;
}
```

BlockGroupContainer.java

```
package structure;

import java.io.Serializable;
import java.util.ArrayList;

public class BlockGroupContainer implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public String key;
    //total, white, nonwhite fraction of HHs with children
    public double tcf, wcf, ncf;
    //overall, white, and nonwhite income group fractions
    public double[] incomeGroup = new double[16];
    public double[] whiteIncome = new double[16];
    public double[] nonwhiteIncome = new double[16];
```

```
    public CensusMoves censusMoves;
    public ArrayList<BlockContainer> blockList = new ArrayList<BlockContainer>();
}
```

BlockContainer.java

```
package structure;

import java.io.Serializable;
import java.util.ArrayList;

public class BlockContainer implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public String key;
    public String bgName;
    public int parcelNum;
    public double owp, rwp, cenX, cenY;
    public ParcelStatus parcelStatus;
    public ArrayList<ParcelContainer> parcelList = new ArrayList<ParcelContainer>();
}
```

ParcelContainer.java

```
package structure;

import java.io.Serializable;

public class ParcelContainer implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public String key;
    public double xPos;
    public double yPos;
    public String bName;
    public String bgName;
    public String pCode;
    public int addressNum;
    public String street;
    public OwnerContainer owner;
}
```

OwnerContainer.java

```
package structure;

import java.io.Serializable;

public class OwnerContainer implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public String key;
    public String pName;
    public String moveCode;
    public String lastName;
    public int lastNameCt;
}
```

ParcelStatus.java

```
package structure;

import java.io.Serializable;

public class ParcelStatus implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public int own;
    public int rent;
    public int vacant;

    public ParcelStatus(int own, int rent, int vacant){
        this.own = own;
        this.rent = rent;
        this.vacant = vacant;
    }
}
```

CensusMoves.java

```
package structure;

import java.io.Serializable;

public class CensusMoves implements Serializable {
    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public double stayed;
    public double movedNear;
    public double movedFar;

    public CensusMoves(double stayed, double movedNear, double movedFar){
        this.stayed = stayed;
        this.movedNear = movedNear;
        this.movedFar = movedFar;
    }
}
```

CaseTown Package

ModelObjects.java

```
package caseTown;

import java.io.Serializable;
import java.util.ArrayList;

public class ModelObjects implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public ArrayList blockGroupList = new ArrayList();
    public ArrayList blockList = new ArrayList();
    public ArrayList parcelList = new ArrayList();
    public ArrayList householdList = new ArrayList();
    public double[][] blockDistance;
    public ArrayList vacantParcelList;

    public ArrayList vacantParcels(){
        vacantParcelList = new ArrayList();
        int size = parcelList.size();
        for (int i=0; i<size; i++){
```

```
            Parcel p = (Parcel)parcelList.get(i);
            if (p.myHousehold==null) vacantParcelList.add(p);
        }
        return vacantParcelList;
    }

}
```

BlockGroup.java

```
package caseTown;

import java.io.Serializable;
import java.util.ArrayList;

public class BlockGroup implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public ArrayList blockList = new ArrayList();
    public ArrayList householdList = new ArrayList();
    public String name;
    public double[] incomeGroup = new double[16];
    public double[] whiteIncome = new double[16];
    public double[] nonwhiteIncome = new double[16];
    public double incAvg, wcf, ncf;

    public double calcBlockGroupIncome(){
        int size = householdList.size();
        incAvg = 0;
        for (int i=0; i < size; i++){
            Household h = (Household)householdList.get(i);
            incAvg += h.income;
        }
        incAvg = incAvg/size;
        return incAvg;
    }
}
```

Block.java

```
package caseTown;

import java.io.Serializable;
import java.util.ArrayList;

import structure.ParcelStatus;

public class Block implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public ArrayList parcelList = new ArrayList();
    public BlockGroup myBlockGroup;
    public String name;
    public ParcelStatus parcelStatus;
    public double xAvg, yAvg, xCen, yCen;
    public int seq;
    public double incAvg;

    public double calcBlockIncome(){
        incAvg = 0;
        int pSize = parcelList.size();
        int hSize = 0;
        for (int i = 0; i<pSize; i++){
            Parcel p = (Parcel)parcelList.get(i);
```

```java
            if (p.myHousehold != null){
                Household h = p.myHousehold;
                incAvg += h.income;
                hSize ++;
            }
        }
        incAvg = incAvg/hSize;
        return incAvg;
    }

}
```

Parcel.java

```java
package caseTown;

import java.io.Serializable;

public class Parcel implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public Block myBlock;
    public String name;
    public Household myHousehold;
    public boolean ownerOccupied; //default is false
    public double xPos;
    public double yPos;
    public double priorOwnerProb;

    public boolean priorOwnerOccupied; //for rent-own transition

    public boolean setPrior(){
        double toss = Math.random();
        if (toss>priorOwnerProb) priorOwnerOccupied = true;
        return priorOwnerOccupied;
    }
}
```

Household.java

```java
package caseTown;

import java.io.Serializable;

public class Household implements Serializable {

    /**
     *
     */
    private static final long serialVersionUID = 1L;
    public Parcel myParcel;
    public String name;
    public double income;
    public boolean white;
    public boolean children;
}
```

Appendix G: AnyLogic Model Structure

Main Active Object Class

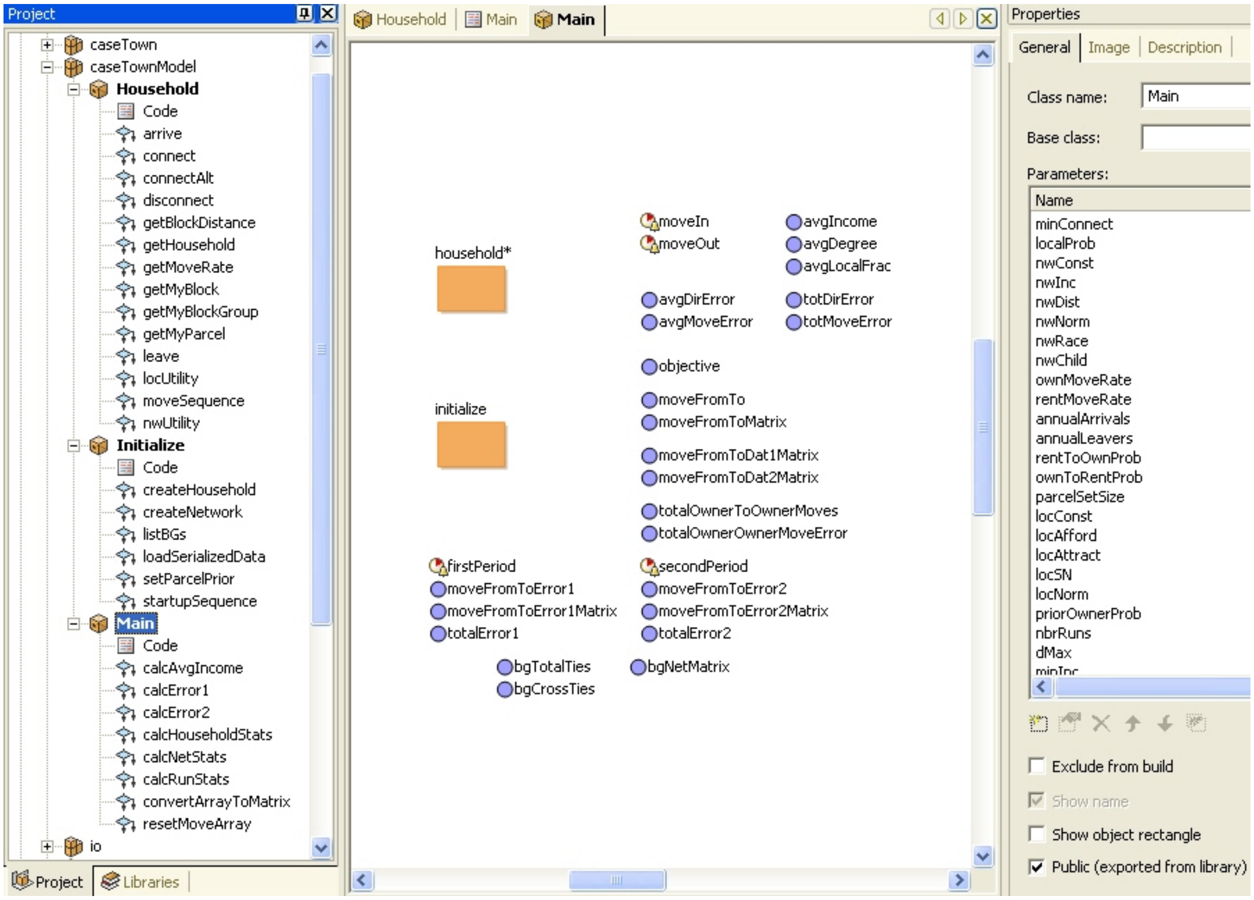

Figure 37. Screenshot of Main Active Object Class Structure

Main Active Object Parameters

```
integer minConnect = 3;
real localProb, nwConst, nwInc, nwDist, nwNorm, nwRace, nwChild //estimated
real ownMoveRate = 0.1;
real rentMoveRate = 10*ownMoveRate;
real annualArrivals = 500;
real annualLeavers = annualArrivals;
real rentToOwnProb = 0.5;
real ownToRentProb = 0.5;
integer parcelSetSize = 15;
real locConst, locAfford, locSN, locNorm; //estimated
real priorOwnerProb = 0.5;
integer nbrRuns = 25;
real dMax = 34736.97; //maximum block distance in feet
real minInc = 5;
real maxInc = 250;
```

Main Active Object Variables

```
real avgIncome, avgDegree, localFrac;
integer totDirError, totMoveError;
real avgDirError, avgMoveError;
real objective;
integer totalOwnerToOwnerMoves, totalOwnerMoveError,
```

```
int[][] moveFromTo = new int[29][29];
Matrix(29,29) moveFromToMatrix, moveFromToDat1Matrix, moveFromToDat2Matrix;
int[][] moveFromToError1 = new int[29][29];
int[][] moveFromToError2 = new int[29][29];
Matrix(29,29) moveFromToError1Matrix, moveFromToError2Matrix;
integer totalError1, totalError2;
int[] bgTotalTies = new int[28];
int[][] bgCrossTies = new int[28][28];
Matrix(28,28) bgNetMatrix;
```

Main Active Object Code

Startup code

```
initialize.startupSequence(); //call the startup sequence
```

Additional class code

```
caseTown.ModelObjects mo = (caseTown.ModelObjects) initialize.mo;
final static int[][] moveFromToDat1 = {
{0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,12},
{0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,5,0,0,0,0,0,0,0,1,1,0,0,0,86},
{0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,32},
{0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,20},
{0,1,0,0,2,0,0,0,0,4,0,0,0,0,0,6,0,1,0,0,0,0,0,0,1,1,1,0,117},
{0,0,0,0,1,0,0,0,0,2,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,1,0,1,52},
{0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,26},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,11},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,20},
{1,4,0,0,5,0,0,0,0,6,0,0,0,0,0,8,0,0,0,0,0,0,0,0,0,0,0,0,140},
{0,0,0,0,1,0,0,0,0,1,0,1,0,0,0,0,0,0,1,1,0,0,0,1,0,0,0,90},
{0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,1,0,50},
{0,3,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,38},
{0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,23},
{0,1,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,41},
{0,4,0,0,0,0,0,0,0,0,0,0,0,0,0,7,0,0,0,0,0,0,0,0,0,0,0,0,133},
{0,3,0,0,0,0,0,0,1,1,1,0,1,0,0,4,1,0,0,0,0,0,0,0,0,0,0,0,81},
{0,2,0,0,1,0,0,0,0,1,0,0,0,0,0,3,0,1,0,0,0,0,0,0,0,0,0,0,66},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,7},
{0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,53},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,7},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,10},
{0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,35},
{0,1,0,0,1,0,0,0,1,0,0,0,1,0,1,0,0,1,0,0,0,0,0,1,0,0,1,0,87},
{0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,36},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,51},
{0,0,0,0,2,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,93},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,12},
{23,79,20,22,134,58,28,2,19,150,91,46,44,21,45,123,94,79,7,62,5,15,52,80,36,49,79,9,0}
};

final static int[][] moveFromToDat2 = {
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,6},
{0,2,0,0,1,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,0,54},
{0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,33},
{0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,15},
{0,2,0,0,3,0,0,0,0,0,0,0,0,0,0,5,0,1,0,0,0,0,0,0,0,0,0,0,97},
{0,0,0,0,0,2,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,51},
{0,0,0,0,0,1,2,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,30},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,10},
{0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,12},
{0,1,0,0,1,0,1,0,0,2,0,0,0,0,0,5,2,0,0,0,0,0,0,0,0,0,0,0,110},
{1,0,0,0,1,0,1,0,0,1,0,0,0,0,0,1,0,1,1,0,0,0,0,1,2,0,0,0,58},
{0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,33},
{0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,27},
{0,0,0,0,0,0,0,0,0,0,1,0,0,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,20},
{0,1,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,30},
{0,2,0,0,1,0,0,0,1,0,3,0,0,0,0,6,0,0,0,0,0,0,0,0,0,0,0,0,92},
{0,0,1,0,3,0,0,0,0,3,0,0,0,0,0,3,1,0,0,0,0,0,0,0,0,0,0,0,60},
{0,2,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,38},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,12},
{0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,38},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,3},
```

```
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,10},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,28},
{0,2,0,0,2,0,1,0,0,0,0,0,0,0,0,2,1,0,0,0,0,0,0,1,0,0,0,0,56},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,1,0,0,0,0,0,0,0,0,21},
{0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,37},
{0,2,0,0,2,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,2,0,69},
{0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,13},
{19,51,20,9,98,44,20,9,15,129,59,23,22,13,25,117,64,45,6,37,4,10,25,44,17,39,70,7,0}
};

public boolean needRepeat() {
  calcRunStats();
  return Engine.getReplication() < nbrRuns;}
```

Main Active Object Algorithmic Functions

real calcAvgIncome()

```
double income = 0;
double avgIncome;
int numH = household.size();
for (int i=0; i<numH; i++){
  Household h = (Household)household.get(i);
  income += h.getHousehold().income;
}
avgIncome = income/numH;
return avgIncome;
```

void calcError1()

```
totalError1 = 0;
totalOwnerToOwnerMoves = 0;
for(int i = 0; i < 28; i++){
for(int j = 0; j < 28; j++){
moveFromToError1[i][j] = (int) abs(moveFromTo[i][j] - moveFromToDat1[i][j]);
totalError1+=moveFromToError1[i][j];
totalOwnerToOwnerMoves += moveFromTo[i][j] ;
}}
totalOwnerOwnerMoveError = (int) abs(totalOwnerToOwnerMoves - 139) ;
int sumErr = totalError1+totalOwnerOwnerMoveError;
int n = (int)Engine.getReplication();
totDirError += totalError1;
totMoveError += totalOwnerOwnerMoveError;
```

void calcNetStats()

```
double totInc, myInc, myLocalFrac, totLocalFrac;
int numH = household.size();
int numF, totF, totLocals, myLocals;
totF = 0; totInc = 0; totLocals = 0; myInc = 0; myLocalFrac = 0; totLocalFrac = 0;
for (int i=0; i<numH; i++){
  Household h = (Household)household.get(i);
  myInc = h.getHousehold().income;
  totInc += myInc;
  caseTown.BlockGroup bg = (caseTown.BlockGroup) h.getMyBlockGroup();
  int m = initialize.mo.blockGroupList.indexOf(bg);
  numF = h.Network.size();
  bgTotalTies[m]+=numF;
  totF += numF;
  myLocals = 0;
  for (int j=0; j<numF; j++){
    Household hj = (Household)h.Network.get(j);
    caseTown.BlockGroup bgj = (caseTown.BlockGroup) hj.getMyBlockGroup();
    int mj = initialize.mo.blockGroupList.indexOf(bgj);
    bgCrossTies[m][mj]+=1;
    if (bg==bgj) myLocals++;
  }
  totLocals += myLocals;
  myLocalFrac = ((double)myLocals)/((double)numF);
  totLocalFrac += myLocalFrac;
}
avgIncome = totInc/numH;
avgDegree = ((double)totF)/numH;
avgLocalFrac = (double) totLocalFrac/numH;
```

```
double cell;
for(int i = 0; i < 28; i++){
  for(int j = 0; j < 28; j++){
    cell = ((double)bgCrossTies[i][j])/((double)bgTotalTies[i]);
    bgNetMatrix.set(i,j,cell);
  }
}
```

void calcRunStats()

```
int n = (int)Engine.getReplication(); //nbrRuns;
avgDirError = (double) totDirError/n;
avgMoveError = (double) totMoveError/n;
objective = avgMoveError + avgDirError;
```

void convertArrayToMatrix()

```
for(int i = 0; i < 29; i++){
  for(int j = 0; j < 29; j++){
    moveFromToDat1Matrix.set(i,j,moveFromToDat1[i][j]);
    moveFromToDat2Matrix.set(i,j,moveFromToDat2[i][j]);
    moveFromToMatrix.set(i,j,moveFromTo[i][j]);
    moveFromToError1Matrix.set(i,j,moveFromToError1[i][j]);
    moveFromToError2Matrix.set(i,j,moveFromToError2[i][j]);
  }
}
```

void resetMoveArray()

```
for(int i = 0; i < 29; i++){
for(int j = 0; j < 29; j++){
moveFromTo[i][j]=0;
}}
```

Main Active Object Timers

moveIn

```
Cyclic Timer
Timeout: exponential(annualArrivals)
Expiry action:
caseTown.ModelObjects mo = (caseTown.ModelObjects) initialize.mo;
int sizeV = mo.vacantParcelList.size();
if (sizeV>0){
  caseTown.Household hj = new caseTown.Household();
  mo.householdList.add(hj);
  hj.income = DistrUniform.sample(minInc,maxInc);
  //find a placeholder parcel at random
  int j = uniform_discr(0, sizeV-1);
  caseTown.Parcel pj = (caseTown.Parcel)mo.vacantParcelList.get(j);
  caseTown.Parcel pMax = pj;
  hj.myParcel = pj;
  Household h = create_household();
  h.stateHousehold = hj;
  h.arrive(pj);
  h.run();
  double toss = DistrUniform.sample(0,1);
    if (toss>priorOwnerProb) h.priorOwner = true;
  double maxUtil = 0;
  sizeV = mo.vacantParcelList.size();
  int sizeRE = (int) min(sizeV, parcelSetSize);
  for (int i=0; i<(sizeRE); i++){
    int startP = h.ParcelSet.size();
    for (int k=0; h.ParcelSet.size()<startP+1; k++){
      j = uniform_discr(0, sizeV-1);
      pj = (caseTown.Parcel)mo.vacantParcelList.get(j);
      if (pj.myHousehold==null && !h.ParcelSet.contains(pj)){
        h.ParcelSet.add(pj);
        double jUtil = h.locUtility(pj);
        if (jUtil > maxUtil) {
          maxUtil = jUtil;
          pMax = pj;
        }//endif
      }//endif
    }//endfor
```

```
  }//endfor
  h.leave(); //empty the placeholder parcel
  h.arrive(pMax);
  caseTown.Parcel p = h.getMyParcel();
  caseTown.BlockGroup bg = h.getMyBlockGroup();
  int m = mo.blockGroupList.indexOf(bg);
  if(p.ownerOccupied==true) moveFromTo[28][m]++;
  for (int n = 0; n < minConnect; n++){
    h.connect();
  }
  calcAvgIncome();
}
```

moveOut

```
Cyclic Timer
Timeout: exponential(annualLeavers)
Expiry action:
caseTown.ModelObjects mo = (caseTown.ModelObjects) initialize.mo;
int numH = household.size();
if (numH>0){
  int j = uniform_discr(0,household.size()-1);
  Household h = (Household)household.item(j);
  caseTown.BlockGroup bg = h.getMyBlockGroup();
  int m = mo.blockGroupList.indexOf(bg);
  for (int i=0; i<h.Network.size(); i++){
    Household hi = (Household) h.Network.get(i);
    hi.Network.remove(h);
    //household is removed from network, not vice versa
  }
  h.leave();
  if(h.priorOwner==true) moveFromTo[m][28]++;
  household.remove(h);
  dispose_household(h);
  //household is NOT removed from the java ArrayList
  calcAvgIncome();
}
```

firstPeriod

```
Expire Once
Timeout: 2.0
calcError1();
convertArrayToMatrix();
resetMoveArray();
```

Initialize Active Object Class

Initialize Active Object Code

Additional class code

```
Main main = (Main)(Engine.getRoot());

public static caseTown.ModelObjects mo;

public static final String filename =
"C:\\Saradocs\\PhD\\Model\\SaraWorking\\modelObjectsFinal.ser";
```

Initialize Active Object Algorithmic Functions

void startupSequence()

```
loadSerializedData( filename );
//filename is defined in Initialize additional class code

createHousehold();
//create households from the imported list, not autocreate

setParcelPrior();
createNetwork();
main.calcNetStats();
mo.vacantParcels();
```

```
void loadSerializedData()
mo = (caseTown.ModelObjects)io.Serializer.load(filename);

void createHousehold()
int nbrHH = mo.householdList.size();
for(int i=0; i<nbrHH; i++){
  Household h = main.create_household();
  h.stateHousehold = (caseTown.Household) mo.householdList.get(i);
  h.run();
}

void setParcelPrior()
int nbrP = mo.parcelList.size();
for(int i=0; i<nbrP; i++){
  caseTown.Parcel p = (caseTown.Parcel) mo.parcelList.get(i);
  p.priorOwnerProb = main.priorOwnerProb;
  p.setPrior();
}

void createNetwork()
int nbrH = main.household.size();
for (int i = 0; i < nbrH; i++){
  Household hi = main.household.item(i);
  for (int j = 0; j < main.minConnect; j++){
    hi.connect();
   }
}
```

Household Active Object Class

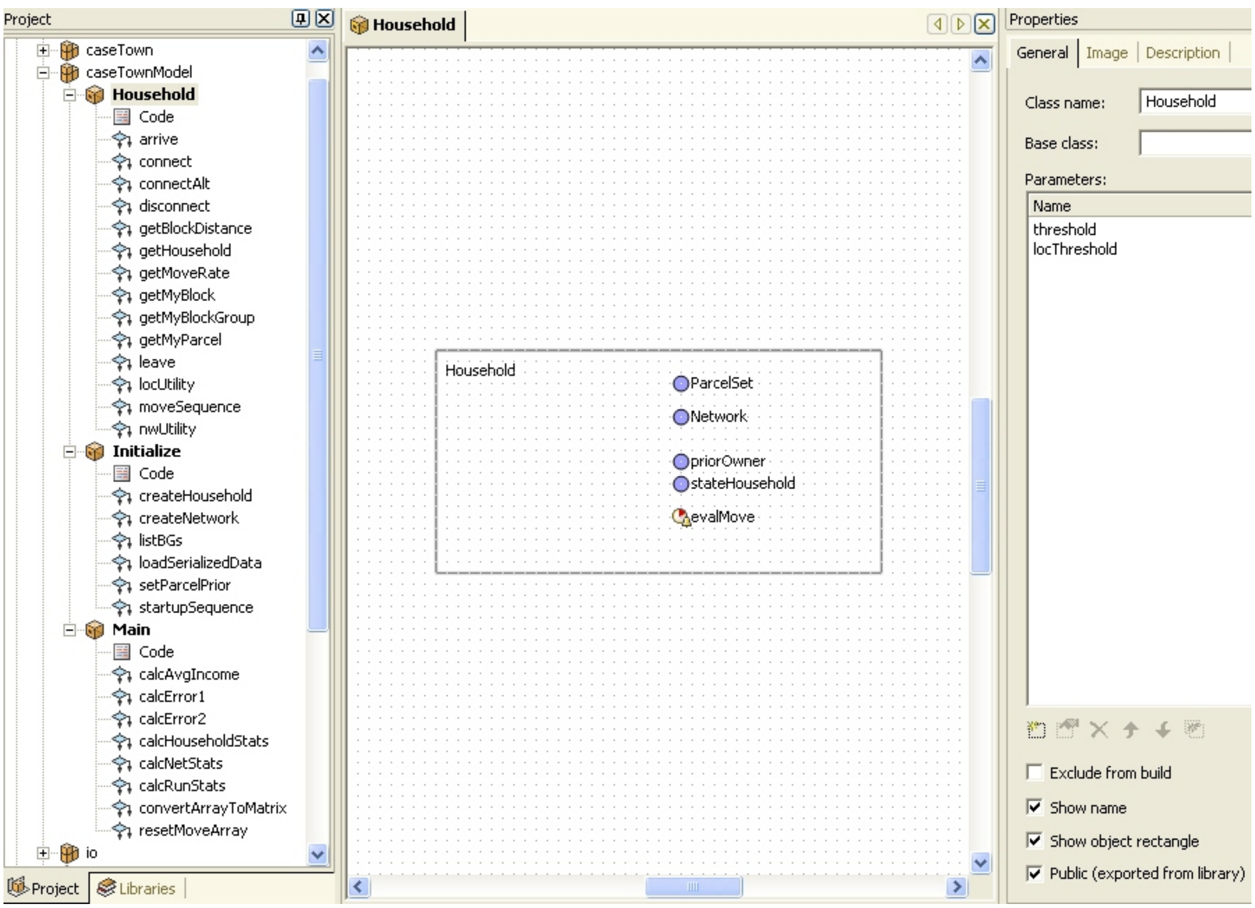

Figure 38. Screenshot of Household Active Object Class Structure

Household Active Object Parameters

```
real threshold = DistrUniform.sample(0,1); //idiosyncrasy in criteria for network connection
real locThreshold = DistrUniform.sample(0,1); //idiosyncrasy in current location satisfaction
```

Household Active Object Variables

```
Vector ParcelSet = new Vector();
Vector Network = new Vector();
boolean priorOwner;
caseTown.Household stateHousehold;
```

Household Active Object Code

Additional class code

```
Main main = (Main)(Engine.getRoot());
caseTown.ModelObjects mo = (caseTown.ModelObjects) main.initialize.mo;
```

Household Active Object Algorithmic Functions

caseTown.Household getHousehold()

```
return (caseTown.Household) stateHousehold; //assigned in initialize.createHousehold()
```

caseTown.Parcel getMyParcel()

```
return getHousehold().myParcel;
```

caseTown.Block getMyBlock()

```
return getMyParcel().myBlock;
```

caseTown.BlockGroup getMyBlockGroup()

```
return getMyBlock().myBlockGroup;
```

void moveSequence(caseTown.Parcel pNew)

```
caseTown.BlockGroup bg = getMyBlockGroup();
int i = mo.blockGroupList.indexOf(bg);
leave();
arrive(pNew);
caseTown.Parcel p = getMyParcel();
bg = getMyBlockGroup();
int j = mo.blockGroupList.indexOf(bg);
if(priorOwner==true && p.ownerOccupied==true){
  main.moveFromTo[i][j] ++;
}//end owner moving
if(priorOwner==true && p.ownerOccupied==false){
  main.moveFromTo[i][28] ++; //j of 28 is leaving owner status
}//end owner leaving
if(priorOwner==false && p.ownerOccupied==true){
  main.moveFromTo[28][j] ++; //i of 28 is new owner status
}//end owner arriving
connect(); //definitely adds a tie
disconnect();//does not necessarily remove a tie
```

void leave()

```
caseTown.Parcel p = getMyParcel();
caseTown.BlockGroup bg = getMyBlockGroup();
p.myHousehold = null;
mo.vacantParcelList.add(p);
if (p.ownerOccupied == true) {
  p.priorOwnerOccupied = true;
  p.ownerOccupied = false;
  priorOwner = true;
}
caseTown.Household h = getHousehold();
h.myParcel = null;
bg.householdList.remove(h);
```

void arrive(caseTown.Parcel p)

```
caseTown.Household h = getHousehold();
p.myHousehold = h;
h.myParcel = p;
mo.vacantParcelList.remove(p);
```

```
caseTown.BlockGroup bg = getMyBlockGroup();
bg.householdList.add(h);
double ro = main.rentToOwnProb;
double or = main.ownToRentProb;

//rent-own prior status
boolean ho = priorOwner;
boolean po = p.priorOwnerOccupied;

if((ho==true)&&(po==true)){
  p.ownerOccupied = true;
}
else if((ho==false)&&(po==true)){
  double toss = DistrUniform.sample();
  if(toss>ro) p.ownerOccupied = true;
}
else if((ho==true)&&(po==false)){
  double toss = DistrUniform.sample();
  if(toss>or) p.ownerOccupied = true;
}
```

void connect()

```
Household hi = Household.this;
double util, p, t, r;
double lp = main.localProb;
double br = DistrUniform.sample(0,1);
int startN = Network.size();
int j;
caseTown.BlockGroup bg = (caseTown.BlockGroup)getMyBlockGroup();

//the for condition enables just ONE new connection
for (int f=0; Network.size() < (startN+1); f++){
    j = uniform_discr(0, main.household.size()-1);
    Household hj = (Household) main.household.item(j);
    caseTown.BlockGroup bgj = (caseTown.BlockGroup) hj.getMyBlockGroup();
    if (((lp > br)&&(bg == bgj))||((lp<br)&&(bg!=bgj))) {
      util = nwUtility(hj);
      p = Math.exp(util)/(1+Math.exp(util));
      r = DistrUniform.sample(0,1); //time of day
      t = r; //*Math.max(threshold, hj.threshold);//ensure mutual preference
      if (p > t && !Network.contains(hj)){
        hi.Network.add(hj);
        hj.Network.add(hi);
      }
    }
}
```

void disconnect()

```
Household hi = Household.this;
double util, p, t;
int jMin = 0;
Household hMin = (Household) Network.get(0);
double minUtil = 0; //initialization only;
int startN = Network.size();
for (int j=0; j<startN; j++){
  Household hj = (Household) Network.get(j);
  util = nwUtility(hj);
  if (j==0) minUtil = util;
  if (minUtil<util) {
    minUtil = util;
    jMin = j;
    hMin = hj;
  }
}
p = 0; //Math.exp(minUtil)/(1+Math.exp(minUtil));
t = Math.max(threshold, hMin.threshold);
if (p < t){
  hi.Network.remove(hMin);
  hMin.Network.remove(hi);
}
```

```
real nwUtility(Household hi)
Household hi = Household.this;
double ii = hi.getHousehold().income;
double C = main.nwConst;
double ad = main.nwDist;
double ai = main.nwInc;
double ar = main.nwRace;
double ak = main.nwChild;
double b = main.nwNorm;
double dMax = main.dMax;
double ij, i, d, r, k, util;
boolean wi = hi.getHousehold().white;
boolean wj = hj.getHousehold().white;
boolean ki = hi.getHousehold().children;
boolean kj = hj.getHousehold().children;

if (wi==wj) r=1; else r=0;
if ((ki==kj)&&(ki==true)) k=1; else k=0;

d = getBlockDistance(hj)/dMax; //divide by max blockdist (ft)
ij = hj.getHousehold().income;
i = Math.abs(ii-ij)/245; //245 is max-min income
util = (C - ad*d - ai*i + ar*r + ak*k)/b;

return util;

real locUtility(caseTown.Parcel p)
double hi, bi, ai;
double afford, attract, netFrac, u;
double wAff, wAtt, wSN, C, normB;

wAff = main.locAfford;
wAtt = main.locAttract;
wSN = main.locSN;
C = main.locConst;
normB = main.locNorm;

hi = getHousehold().income;

caseTown.Block b = (caseTown.Block)p.myBlock;
bi = b.calcBlockIncome();
ai = main.avgIncome;

if (p == getMyParcel()) afford = 0;
else {
  afford = (hi-bi)/bi;
  afford = min(0, afford); //only if bi > hi
}
attract = (bi-ai)/ai;

caseTown.BlockGroup bg = (caseTown.BlockGroup)b.myBlockGroup;

int n = Network.size();
int bn = 0;
for (int j=0; j<n; j++){
  Household hj = (Household)Network.get(j);
  caseTown.BlockGroup bgj = (caseTown.BlockGroup)hj.getMyBlockGroup();
  if (bg == bgj) bn++;
}
if (n==0) netFrac = 0;
else netFrac = bn/n;
u = (C + wAff*afford + wAtt*attract + wSN*netFrac)/normB;

return u;

real getMoveRate()
caseTown.Household h = getHousehold();
caseTown.Parcel p = getMyParcel();
boolean po = p.ownerOccupied;
if (po == true) return main.ownMoveRate;
else return main.rentMoveRate;
```

Household Active Object Timer

```
evalMove
Cyclic Timer
Timeout: exponential(getMoveRate())
Expiry Action:
caseTown.Parcel p = getMyParcel();
double myUtil = locUtility(p);
double prob;
prob = exp(myUtil)/(1+exp(myUtil));

int sizeV = mo.vacantParcelList.size();
if (prob < locThreshold && sizeV >0){
  double maxUtil = myUtil;
  caseTown.Parcel pMax = p;
  ParcelSet.clear();
  int sizeRE = (int) min(sizeV,main.parcelSetSize);
  for (int i=0; i<sizeRE; i++){
    int startP = ParcelSet.size();
    int j;
    for (int k=0; ParcelSet.size()<startP+1; k++){
      j = uniform_discr(0, sizeV-1);
      caseTown.Parcel pj = (caseTown.Parcel)mo.vacantParcelList.get(j);
      if (pj.myHousehold==null && !ParcelSet.contains(pj)){
        ParcelSet.add(pj);
        double jUtil = locUtility(pj);
        if (jUtil > maxUtil) {
          maxUtil = jUtil;
          pMax = pj;
        }//endif
      }//endif
    }//endfor
  }//endfor
  if (pMax != p){
    moveSequence(pMax);
  }//endif
  ParcelSet.clear();
}//end unhappy condition
```

Appendix H: System Dynamics Basics

This appendix outlines some of the basics of system dynamics. The "system" part of system dynamics implies a consideration of closed-loop systems, in which feedback mechanisms are of utmost importance. The first type of feedback mechanism is reinforcing feedback (the "vicious" or "virtuous" cycle), implying sustained growth or decline. An example is the learning curve illustrated in Figure 39 below: as cumulative production increases, unit costs decline, and prices decline accordingly (the + indicates change in the same direction as the preceding variable). Price is inversely related to market share, so as it declines, the market share (demand) will increase. As demand increases, then cumulative production increases. The key here is to multiply the negative(-) connections; if there are an even number of negative connections, the feedback is positive or reinforcing because the connections cancel each other through multiplication. If there are an odd number of negative connections, the feedback is negative or balancing.

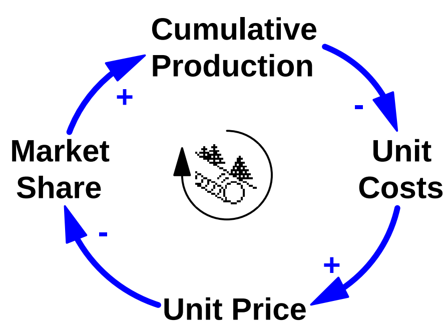

Figure 39. Example of Reinforcing Feedback Loop

The second type of feedback mechanism is known as a balancing feedback loop. The balancing loop exhibits "goal-seeking" behavior: rather than growth or decline, the balancing feedback seeks to attain equilibrium level. The presence of delays in the system (illustrated by hatchet marks on the arrows) can result in oscillation around the desired state. An example of the reinforcing feedback mechanism could be efforts to mitigate pollution, as illustrated in Figure 40 below. Suppose there is a desired pollution level that serves as the goal. The difference between this goal and the actual level of pollution results in a pollution gap. The greater the gap, the greater the concern for pollution. And the greater the concern, the greater the action to reduce pollution. This of course takes time, and so the delay is represented by a hatchet, but over that delay the actual pollution level declines. This decline means the gap is smaller so concerns lessen, so actions lessen, and so on.

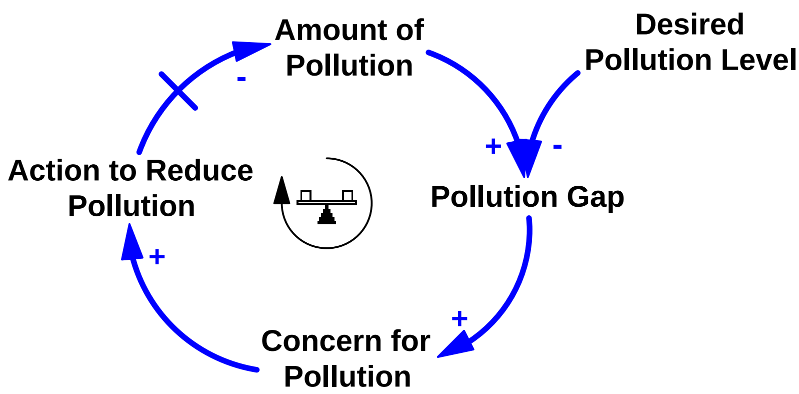

Figure 40. Example of Balancing Feedback Loop

Real-world systems have combinations of balancing and reinforcing loops. You would not see one type of feedback loop acting in isolation. S-shaped growth occurs when reinforcing and balancing loops are present: there is a period of sustained growth followed by a damping of that growth as the system seeks its equilibrium state. Figure 41 illustrates interacting reinforcing and balancing effects.

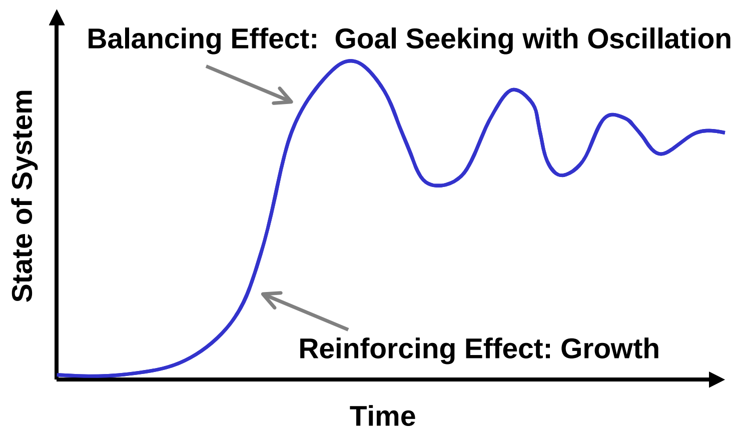

Figure 41. Example of Interaction of Reinforcing and Balancing Effects

If feedback loops are the "system," then stocks and flows are the "dynamics" of system dynamics. The presence of a stock and flow structure means that the system can contain inertia, memory, or delays. The stocks are at the heart of this. They represent accumulation of something (material, energy, or information) and the levels of the stocks characterize the state of the system. Flows represent the rates that enter or leave a stock, and thus represent how the stock changes over time. In a system "snapshot", the flows would be invisible, and the stocks would be apparent.

The classic example is the bathtub, as illustrated in Figure 42 below. There are two flows: inflow of water from the faucet (controlled by an inflow valve), and flow out of the bathtub down the drain (controlled by an outflow valve). The water in the bathtub is the stock, perhaps quantified in terms of gallons, while the rates of flow would be quantified in terms of gallons per minute. (In mathematical terms, the stock level is determined by integrating the rates of flow.)

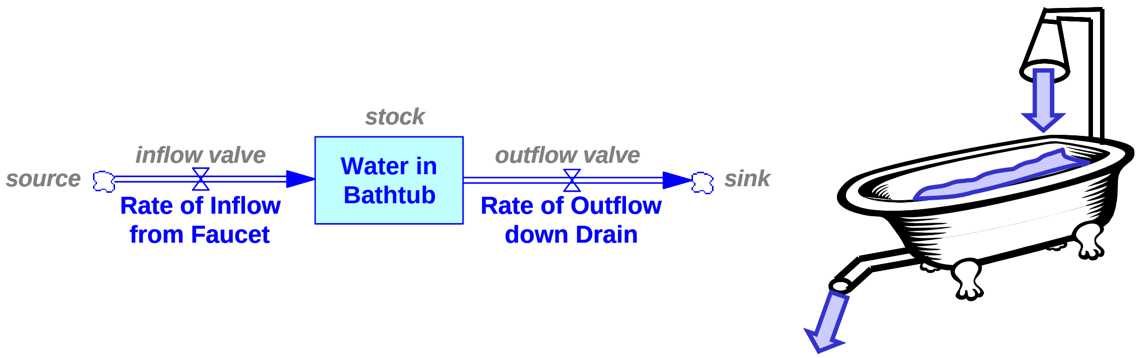

Figure 42. Stock and Flow Representation of Bathtub Dynamics

Stocks and flows combine with feedback loops to create dynamic systems. Stocks provide information about the system (so the causal arrow would exit a stock), and the stock can only be changed through its flows (so the causal arrows lead to flows, not stocks).

The flow arrows implicitly represent causal arrows leading to the stock variable—for instance, in a causal loop structure, an increase in the rate of outflow would lead to a decrease in the stock and thus be represented by a negative connection.

Excellent introductions to system dynamics are available that also demonstrate the user-friendly Vensim (Sterman 2000) and STELLA (Hannon and Ruth 2001) icon-based software. In advocating the above system representation, the father of system dynamics has observed that "nature only integrates" (Forrester 1996).[22]

[22] The more complete context for Forrester (1996, p. 27)'s assertion is as follows: "One might ask how it is possible to teach behavior of complex dynamic systems in K-12 when the subject has usually been reserved for college and graduate schools. The answer lies in having realized that the mathematics of differential equations has been standing in the way." ... "Differential equations are difficult, weak, confusing, and unrealistic. They often mislead students as to the nature of systems. Mathematicians have had difficulty defining a derivative and there is a reason. Derivatives do not exist except in a mathematician's imagination. Nowhere in nature does nature take a derivative. Nature only integrates, that is, accumulates. Casting behavior in terms of differential equations leaves many students with an ambiguous or even reversed sense of the direction of causality. I have had MIT students argue that water flows out of the faucet because the level of water in the glass is rising; that seems natural to them if the flow has been defined as the derivative of the water level in the glass."