

Chapter 6

Simple Agents, Complex Emergent City: Agent-Based Modeling of Intraurban Migration

Shipeng Sun and Steven M. Manson

Abstract Intraurban migration—residential movement within a metropolitan area—defines the nature of urbanization. Housing location decision making is a complex process driven by the interactions between the housing market and home searchers. Researchers have paid much attention to the environmental, socioeconomic, cultural, and policy features of housing markets. In contrast, housing search has been relatively neglected due to challenges of theory, methodology, and data. This article addresses these challenges by presenting an agent-based model of intraurban migration featuring straightforward and empirically specified rules for housing search. This model is calibrated and validated against real-world housing vacancies and relocation origin–destination pairs extracted from parcel records for the Twin Cities of Minnesota, USA, for 2005–2007. Drawing on these unique data sidesteps a long-standing issue, the prohibitive costs of identifying, recording, and quantifying housing search activities for an entire metropolitan region. Conceptually, this model updates geographic theories of intraurban migration that focus on intervening opportunities and spatial bias. It also methodologically advances the agent-based modeling of urbanization with a high-resolution, empirically specified model that demonstrates how urban pattern emerges from simple rules and interactions. Overall, the model demonstrates that relatively straightforward housing search rules can simulate realistic patterns of intraurban migration.

Keywords Intraurban migration • Agent-based modeling • Housing search • Housing locational decisions

S. Sun (✉)

Department of Environmental Studies, University of Illinois Springfield, One University Plaza,
Springfield, IL 62703, USA

e-mail: sunsp.gis@gmail.com; ssun32@uis.edu

S.M. Manson

Department of Geography, Environment and Society, University of Minnesota, 267 19th Avenue
South, Minneapolis, MN 55455, USA

e-mail: manson@umn.edu

6.1 Introduction

Intraurban migration, or residential moves within a metropolitan area, is a complex process involving the interaction of housing market characteristics with the perceptions of home searchers. Intraurban migration research has a long and rich history in geography and other social sciences, engineering fields, and policy disciplines (Dorigo and Tobler 1983; Clark 1986; Brown and Moore 1970; Roseman 1971; Simmons 1968; Simpson et al. 2008; Clark 2008). There remain significant challenges in terms of data, method, and theory in understanding this form of migration. Data on specific individuals who drive intraurban migration is difficult to obtain and use. Methodologically, there is a need to combine the most common approach, statistical methods, with fast-emerging simulation modeling methods. In terms of theory, the large number of competing explanations for intraurban migration points to the need for continuing work on foundational research on individual behavior. When these challenges of data, method, and theory are taken together, they indicate the need for empirically based approaches that combine statistical and simulation models to develop and test straightforward frameworks for understanding how individual behavior gives rise to aggregate patterns and processes of intraurban migration.

We developed an agent-based model that draws on novel data derived from land parcels to develop and test an updated form of the intervening opportunity theory of intraurban migration. This work is significant in several respects. This conceptual model brings together underexamined geographical and sociological findings to develop and test straightforward spatial behavioral rules that capture key features of intraurban migration. To specify and test this model we developed a new data source, individual migration chains extracted from tax parcel data that allowed us to track the movements of actual households in space and over time. We bring these data and the conceptual model together in an agent-based model that is calibrated and validated with mathematical and statistical approaches, leveraging the relative strengths of these different methods. More broadly, this work addresses the need for simple and generalizable to complement the large and growing body of work that focuses on representing complicated dynamics with extensive and detailed datasets (Brown et al. 2008; Torrens 2012). It also contributes to the fast expanding body of work seeking to simplify complex urban dynamics by using new data sources to develop relatively straightforward and generalized models that capture significant features of urban form and processes (Batty 2008, 2012).

The rest of this paper examines this confluence of data, method, and theory. The next section reviews locational decision-making theories of intraurban migration and proposes a model for housing location decisions with two different strategies. Section 6.3 applies this model to intraurban migration of homeowners in the Twin Cities Metropolitan Area of the USA (TCMA), along the way introducing the use of land parcel data to calibrate and validate an agent-based model of individual migration. Section 6.4 presents the model results, including model validation. The paper concludes with discussion of our findings and their implications for urban agent-based modeling and our understanding of intraurban migration generally.

6.2 Conceptual Model of Intraurban Migration

A core conceptual challenge in understanding intraurban migration is developing theories of how individual behavior leads to complex urban patterns and processes. Intraurban migration has three interrelated components, as introduced by Wolpert (1965) and expanded over the years: (1) housing conditions, or the broad social, demographic, economic, and environmental conditions that trigger household migration; (2) housing utilities, expressed as the balance between utilities of the current housing and expected utilities of other housing opportunities; and (3) housing search, or the search process and perceptions of housing by potential buyers. Intraurban migration research focuses primarily on the first and second components, while a smaller body of work centers on the third component of housing search as a sociospatial process that guides the first and second.

We advance this third component by testing a modified intervening opportunity theory, drawing on sociospatial conceptions of housing perception to examine simple decision-making rules that lead to realistic complex migration patterns in aggregate. This third component helps guide the first two and can be examined separately, which does not minimize the fact that the residential choices of households are driven in part by a host of demographic and socioeconomic characteristics of migrants combined with housing utilities, including housing structure, the biophysical environment, neighborhood quality, as well as accessibility to services (Adams 1984; Quigley and Weinberg 1977; Clark 2008). Additional considerations include household factors ranging from income and race to environmental preferences (Choldin 1973; Pellegrini and Fotheringham 2002; Jones et al. 2004). These factors in turn modify the effect of housing conditions and their attendant perceived difference in utilities (De Jong and Roempke Graefe 2008; Geist and McManus 2008; Mulder 2007; Cooke 2008), interactions with commuting and transportation infrastructure (Clark, Huang, and Withers 2003; Rouwendal and Rietveld 1994), government policies and developer decisions (Brown and Chung 2008), and the lending practices of financial institutions (Brown and Longbrake 1970; Brown and Moore 1970; Clark 1982). In sum, a wide variety of factors influence housing conditions and housing utilities and, by extension, intraurban migration and urbanization more generally.

Despite the importance of housing conditions and housing utilities, the third component of housing search and perception has much to do with the nature of intraurban migration because it guides the effects of the first two components. This component has a distinguished research history, but overall it has received far less attention than the first two components. Theories of housing search complement our understanding of these other components because they posit that there are fundamental regularities in how households perceive housing opportunities. Much of this research emphasizes the distinct spatial and temporal limits of homebuyers' search strategies in local and regional housing markets (Clark 1982; Clark and Flowerdew 1982; Smith et al. 1979). Related work examines how home search and job search interact, particularly in how people think about where they want to live as a function of where they want to work (Waddell 1993; van Ommeren et al. 1997;

cf. Clark and Withers 1999). This research points to the importance of incomplete information and bounded rationality of decision making in intraurban migration, as much of the work on the first two components of housing condition and utility assumes that households have complete information during the housing search and will go to any lengths to find the optimal home. Instead, research on search and perception highlights how housing search is often bounded in space and time, whether by homebuyers' greater knowledge of and comfort with local neighborhoods, bounds on how much time households (or their real estate agents) can devote to the housing search, or a willingness to settle for a new house that is good enough instead of being perfect.

The importance of spatial distance between current and potential housing is a unifying theme in much work on housing search, along with direction to a lesser extent. Most of this research relies on the theoretical antecedent of intervening opportunity theory, developed by Stouffer (1940) to describe the relationship between housing opportunities and moving distances within a metropolitan area. Assuming the quantity of vacant housing units is proportional to the distance from a household's current dwelling, Stouffer posited that the number of households that move a given distance has a logarithmic relationship with the housing opportunities located within that distance because people are likely to choose a vacancy near their current dwelling. The related exponential distribution of moving distances has been validated by empirical research on various metropolitan areas (Clark and Burt 1980; Clark et al. 2003; Quigley and Weinberg 1977). The basic form of this negative exponential distribution of move distances is

$$f(d) = \lambda e^{-\lambda d}, \quad \lambda > 0 \text{ and } d > 0 \quad (6.1)$$

where $f(d)$ is the probability of a household relocating by distance of d and λ is a shape parameter. Mathematically, λ is the reciprocal of average d , or in other words $1/\lambda$ is the average move distance. The intervening opportunities model was extended to the case of interurban migration and evolved into the influential gravity model and related family of spatial interaction models (Guldmann 1999; Jayet 1990; Fotheringham 1983; Cochrane 1975; Ruitter 1967; Erlander 2010). Regardless of variant or degree of sophistication, these models retain at their heart a focus on logarithmic or exponential distance decay in space.

The direction people move is an important consideration alongside distance. Adams (1969) argued that spatial search and residential locational behavior are based on a limited mental map or image of the city. Importantly, this image is sectoral in that it comprises a wedge-shaped region centered on the work-home axis. While Adams took the central city is a proxy for work location, the theory was validated using specific workplace data as well (Clark and Burt 1980; Clark et al. 2003). Move directions can be modeled as a von Mises distribution (Gaile and Burt 1980), which is the counterpart of the normal distribution for directional data spanning 0–360°, with a density function of

$$f(x|\mu, \kappa) = \frac{e^{-\kappa \cos(x-\mu)}}{2\pi I_0(\kappa)} \quad (6.2)$$

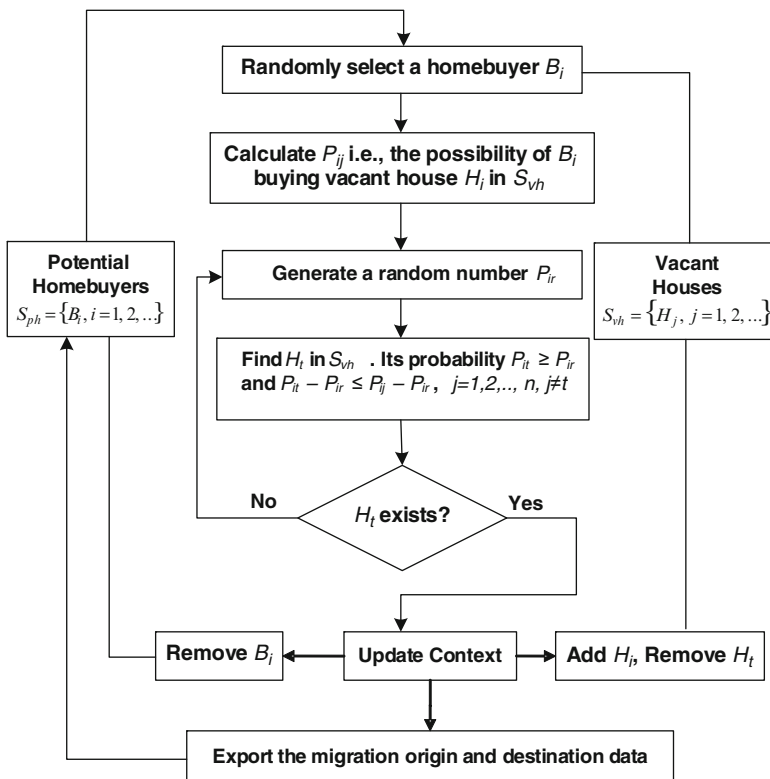


Fig. 6.1 Modeled intraurban migration process

where μ is the mean direction, κ is a measure of variance of directions around μ , and $I_0(\kappa)$ is a modified Bessel function with order zero. When κ is zero, the distribution is uniformly circular (i.e., with equal probability in any direction); when it is larger, the distribution will concentrate around μ in a similar fashion to the normal distribution.

We bring together these sociospatial findings on housing perception via a conceptual agent-based model of the distance and directional relationships among housing vacancies and current dwellings of potential migrants (Fig. 6.1). This model adopts two housing search and relocation strategies—distance-only and distance-plus-direction—that condition intervening opportunity theory with the statistical distribution of moving distances and directions as specified by the negative exponential and von Mises distributions (after Clark et al. 2003). The model is based on two lists for the regional housing market: one of potential homebuyers and another of vacant houses. Each model year, the model iterates over homebuyers randomly. Each homebuyer generates a random number and the vacant house with the closest greater probability is chosen as the destination. When a homebuyer moves, he or she is removed from the buyer list and the vacated house is added into the vacant house list.

The difference between the two homebuyer strategies—distance-only and distance-plus-direction—lies in the probability assigned to vacant houses, P_{ij} , or the possibility of B_i buying vacant house H_j (Fig. 6.1). With the distance-only strategy, each actor agent calculates P_{ij} based on the distance between her current dwelling and a vacant house, where the probability follows a negative exponential distribution. Assuming homebuyer B_i currently lives in H_i , the probability that she chooses house H_j would be

$$P_{ij} = \lambda e^{-\lambda d(H_i, H_j)} \quad (6.3)$$

where λ is a parameter estimated empirically from move distance distribution (more on this below) and $d(H_i, H_j)$ is the distance between H_i and H_j . With the distance-plus-direction strategy, directional bias is also included in the calculation of probability P_{ij} . When relocation is constrained by real housing opportunities, it can be assumed that move direction is independent from move distance (Adams 1969; Clark and Burt 1980; Clark et al. 2003). The von Mises distribution is modeled as two normal distributions with zero and 180° as mean values, respectively:

$$P_{ij} = \lambda e^{-\lambda d(H_i, H_j)} \cdot P_\theta \quad (6.4)$$

where P_θ is the probability that a homebuyer moves in the direction of θ . If we define $\text{Sign}(\theta) = \begin{cases} 1, & \text{if } |\theta| \leq 90 \\ 0, & \text{if } |\theta| > 90 \end{cases}$, then $P_\theta = \text{Sign}(\theta) \cdot N(0, \sigma_1^2) + [1 - \text{Sign}(\theta)] \cdot N(180, \sigma_2^2)$, in which $N(\mu, \sigma^2)$ is normal distribution. When $\text{Sign}(\theta)$ is one, homebuyers move toward the suburbs; when it is zero, they move toward downtown. The standard deviation σ_1 and σ_2 control the extent to which migrant household moves concentrate along the home–downtown corridor. When these deviations are small, houses near the corridor are more likely to be chosen, but when they are large, more houses have greater odds of being chosen. When σ_1 is greater than σ_2 , households are more likely to move to suburbs; when smaller, households move toward downtown.

6.3 Methods

A key methodological challenge in understanding intraurban migration is developing straightforward, empirically specified approaches to model how the choices of individuals generate aggregate migration patterns and processes. The primary form of intraurban migration modeling is mathematical and statistical, ranging from gravity modeling to hedonic specifications to various flavors of (new) economic geography of urban areas. Less common but long-standing is simulation modeling, which has enjoyed renewed interest in the form of agent-based modeling (ABM). We develop an ABM of urban intramigration and use a new form of data to calibrate

this model, namely, empirically specified migration chains from land parcel data. While they have some drawbacks, these data offer several advantages over many other forms of data used to understand the migration choices of individuals.

6.3.1 Data

The paucity of data on the migration choices of individuals remains a critical challenge in understanding intraurban migration. While migration evinces clear patterns such as suburbanization, gentrification, or decline when examined at gross temporal and spatial scales, our understanding of migration at the scales of individuals is limited by the dearth of public data available on movements of individual households in space at the scale of specific housing units and in time at the scale of a year (Adams 1969; Clark 1976, 1986). There are several different ways to garner these data, although we focus on the advantages of parcel data below.

A common approach to measuring intraurban migration is surveying individuals and then reporting on them over large enumeration units. These surveys ask questions about recent moves, such as time since last move or change in commuting time, and range from travel surveys to general instruments such as the American Community Survey (ACS), the American Housing Survey (AHS), the Current Population Survey (CPS), and the Public Use Microdata Samples (PUMS). These surveys are taken of individuals but when reported are aggregated to regions such as census tracts or traffic analysis zones. As a result, these sources offer good information about intraurban migration in general but lack the spatial resolution necessary to analyze individual moves at subregional scales. These data may be downscaled to create statistically plausible individuals (e.g., giving agents an income from a statistical distribution and giving them a random location within a census tract), but this does not link to actual individuals and places (Berger and Schreinemachers 2006). In sum, census-like surveys offer good attribute detail over broad extents at the cost of spatial specificity.

Another common approach to measuring intraurban migration is to gather data on specific households or houses in an area. Directly surveying migrants is a good way to understand their home-seeking behavior, but this approach is expensive and typically reaches only a small subset of migrants. Other sources include city or telephone directories and utility records that can be used to track the moves of individuals from one address to another, although these data are often incomplete and, in cases such as utility records, subject to confidentiality provisions. A related approach is using home sales data to capture attributes of specific houses, but these data usually say little about the search and migration behavior of specific individuals. Overall, data on specific households and houses offer spatial specificity not found in aggregate data noted above, but their use is not without challenges.

We developed a novel form of information on household intraurban migration to address key data challenges, namely, migration chains from land parcel data for an entire region. A migration chain establishes linked pairs of moves, each

defined by a household that leaves a property and one that moves into the just-vacated property. Parcel data are suited to this task when they encompass all home ownership for a specific area; in the Twin Cities, for example, these data describe over one million lots. This research utilizes the annual regional parcel dataset in the TCMA compiled and managed by the regional government, the Metropolitan Council, spanning the seven counties of Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. Relevant information includes owner's name and date last sold; other data vary by jurisdiction, such as square footage of houses and their lots or dwelling type (for a review of these data and those from other locations, see Manson et al. 2009). We identified about 4,800 origin–destination pairs for the years 2005 through 2007, which contain the most complete information for the region and pertain to the period before the US housing market collapsed in 2008.

While developing migration chains from land parcel data is laborious, it can be semiautomated. We developed migration chains for the Twin Cities by comparing the owners of a parcel across years, detecting valid owner changes and matching owners across years. We weeded out transactions, such as speculation and bank sales, that represent ownership change without a household move. We also left out condominiums and apartments given that many are not owner occupied (so renters are not included). We developed software that embodied a multipart strategy to deal with variations and errors in names. All names were uniformly formatted into the order of first name, middle name, and last name. Then an intelligent name comparison routine determined if two different names actually refer to the same person, family, or organization. It employed a dictionary of abbreviations, which records various forms of names for a single institution such as the city of Minneapolis and MPLS and the Minnesota Department of Transportation and MNDOT. It also scanned all parts and letters in two names, and if the percentage of matched parts or letters is beyond a predefined criterion, the two names are defined as the same. For instance, George Washington and G. Washington would be judged as the same person, and George Washington and George and Martha Washington are the same household. We then reviewed all matches manually to minimize dataset errors.

Semiautomated extraction of migration chains from parcel data is not a panacea for migration research, but it offers significant advantages over other approaches. While it can identify housing attributes, such as square feet of number of rooms, it does not provide characteristic of movers, such as age or size of household. It is far more extensive in coverage than most sales databases, much less expensive than surveys of individuals, and provides a level of specificity not found in higher-level data such as the census. As a result, this novel approach to migration data provides critical spatial and temporal information at resolutions sufficient to test theories of individual migration.

6.3.2 Agent-Based Modeling of Migration

We develop an agent-based model of the modified intervening opportunity theory presented above that is calibrated and validated against migration chains derived

from parcel data. The ABM treats residential choice as primarily influenced by distance and direction between movers and vacancies, updating the classic intervening opportunity model with individual agents acting on real-world evidence. Agent-based modeling has garnered a lot of attention for spatially explicit modeling of urbanization and land use more broadly (Gimblett 2002; Parker et al. 2003a; Irwin et al. 2009; see Batty 2008; O'Sullivan 2008). An agent-based model is a computational system composed of semiautonomous software programs (termed agents) that can represent entities ranging from atoms through households to cities. Each agent in the system has its own resources, local context, knowledge, behavioral rules, and goals. Importantly, agents interact with each other and their larger environment. ABMs are increasingly used to understand urban issues such as growth and sprawl, land use and transportation, and racial segregation and residential structure because they explain how simple microbehavior leads to complex macro patterns and processes (Torrens 2006; Fossett 2006; Salvini and Miller 2005; Miller et al. 2004). Using an ABM is important given the intractability of deriving analytical solutions to a system of equations defined by real-world spatial data on thousands of individuals outside of a simplifying mathematical approach or use of a statistical model (Krzanowski and Raper 2001; Kwasnicki 1999). These approaches are commonly used in part because they are powerful, but an ABM, by instantiating in agents the underlying mathematical formulation of intervening opportunity theory, allows exploration of the theory in a real-world context. Marrying mathematical and statistical formalism with agent-based modeling is increasingly seen as a way forward for theoretically derived and empirically tested models of human behavior (Irwin et al. 2009).

The model developed here joins other related efforts that use ABM to understand urban processes. There is a fast-growing body of research that applies this approach to construct models centered on representing the decision-making processes of individuals and their resultant mobility (Haase and Schwarz 2009; Torrens 2012; Kennedy 2012; Parker et al. 2003b; Macy and Willer 2002; An 2012; Matthews et al. 2007; O'Sullivan et al. 2012). These models vary broadly in their degree of specificity and extent to which they are conceptually stylized models. Some attempt to simulate classical urban residential processes and patterns, such as monocentric cities and residential segregation (Benenson and Torrens 2004; Crooks et al. 2008), with highly generalized and stylized models. Others build on these simpler models via greater empirical specification, seeking to simulate urban residential processes including gentrification (Jackson et al. 2008; Diappi and Bolchi 2008; O'Sullivan 2002; Torrens and Nara 2007) and urban sprawl (Brown et al. 2008; Fernandez et al. 2005; Loibl and Toetzer 2003). Other models go even further by offering intricately detailed and data-rich explorations of urban processes underlying complex residential choices within the urban sphere (Birkin and Wu 2012; Zaidi and Rake 2001).

This model is implemented in a spatially explicit agent-based model of land change (Manson and Evans 2007). Agents are software objects, or semiautonomous programs that have their own properties and routines, that exist in an environment composed of raster and vector format layers. Importantly, agents update the

environment by virtue of changing spatial layers after taking actions such as building new houses or moving between houses. The process of modeling TCMA intraurban migration in the model has four steps: (1) establishing the spatiotemporal context, (2) populating agents and environment, (3) running the model to create vacancies and simulate migration, and (4) validating model output (Fig. 6.2).

6.3.2.1 Step 1: Spatiotemporal Context

We model intraurban migration in the seven-county TCMA from 2005 to 2007. As an organizing framework, we adopt standard housing submarkets that map onto well-established neighborhoods as defined by the regional real estate board (see Fig. 6.3). We interpolate the population of each submarket as given by regional government surveys and land-use zoning to fit in these submarkets as a series of raster data layers with a resolution of 100 m.

6.3.2.2 Step 2: Agent Specification

The chief actors are households in the owner-occupied housing sector, housing developers, and governmental institutions. Populating actor agents involves significant simplification, because the core strength of agent-based modeling is illustrating how complex results can arise from simple actions. We focus on three types of agent.

Institutional Agents They shape housing development and migration destination options. The model incorporates the policy effects of the regional planning agency, the Metropolitan Council, and local governments through a set of areas that are off limits to new housing (land reserved for agricultural use or wetland offsets) and areas that are designated for new development (defined by growth zones and sewerage availability). These effects are coded as rules that denote locations where development can and cannot occur.

Developers They expand the housing stock and add new vacancies into the housing market to join the existing vacancies given by the parcel dataset. Developers build new houses that are added to the vacancy lists. Their key characteristic is the rate at which they build houses, which is given empirically by the parcel dataset as 5,392 per year. They build houses at random locations in areas designated by institutional agents. Using developers is a straightforward way of ensuring the growth in housing mirrors that in reality while maintaining an analog to the real world, but their decision making is far simpler than that of real developers.

Households These are the primary agent of interest. Agents are placed in the study area via a polygon file where the number of households within each spatial unit is determined by the actual household population in a given neighborhood, listed in the population data noted above. The migration rate is around 7 % per year for owner-occupied housing and this number of households is placed. Households are assigned

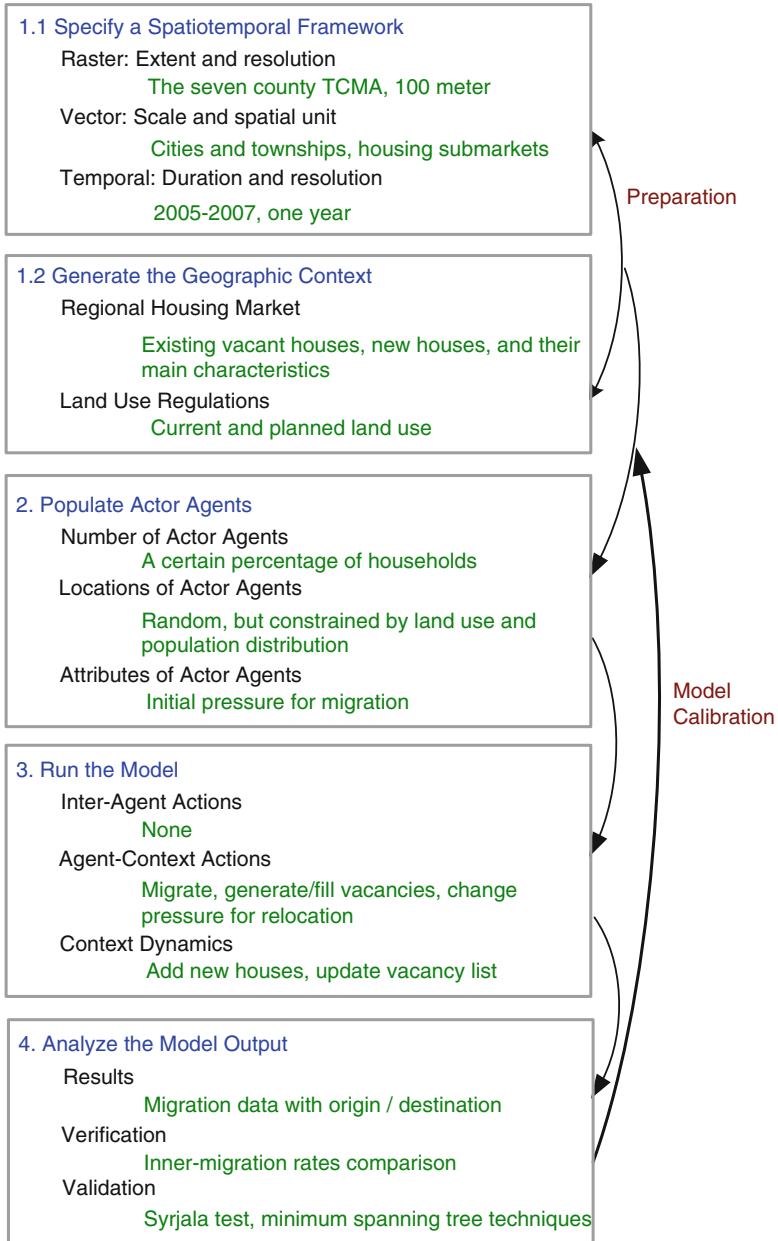


Fig. 6.2 Main modeling steps

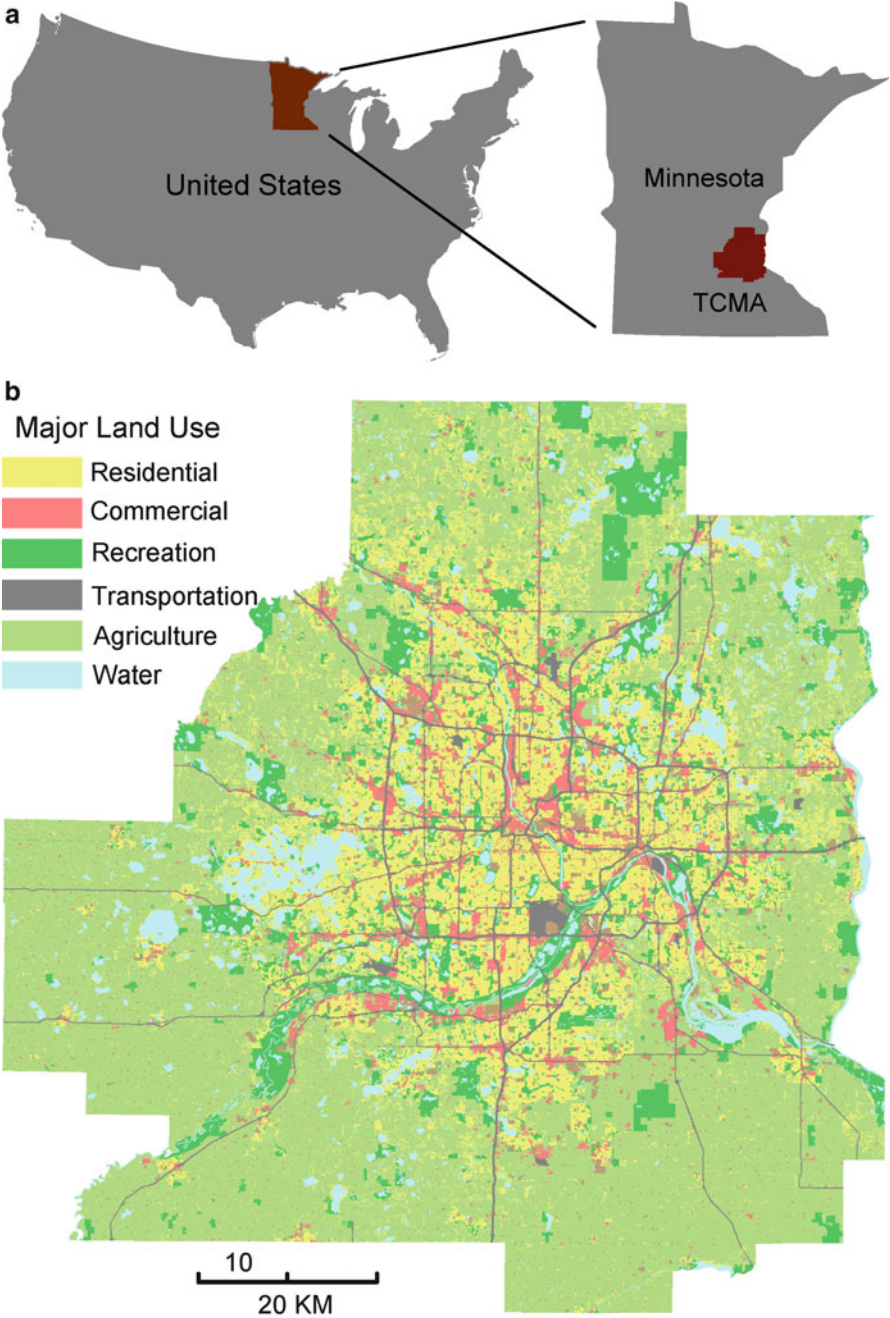


Fig. 6.3 The spatial context of the model (a) modeled area, (b) land-use pattern, (c) spatial configure based on housing submarkets used by realtors, and (d) detailed parcel map in the city of Victoria

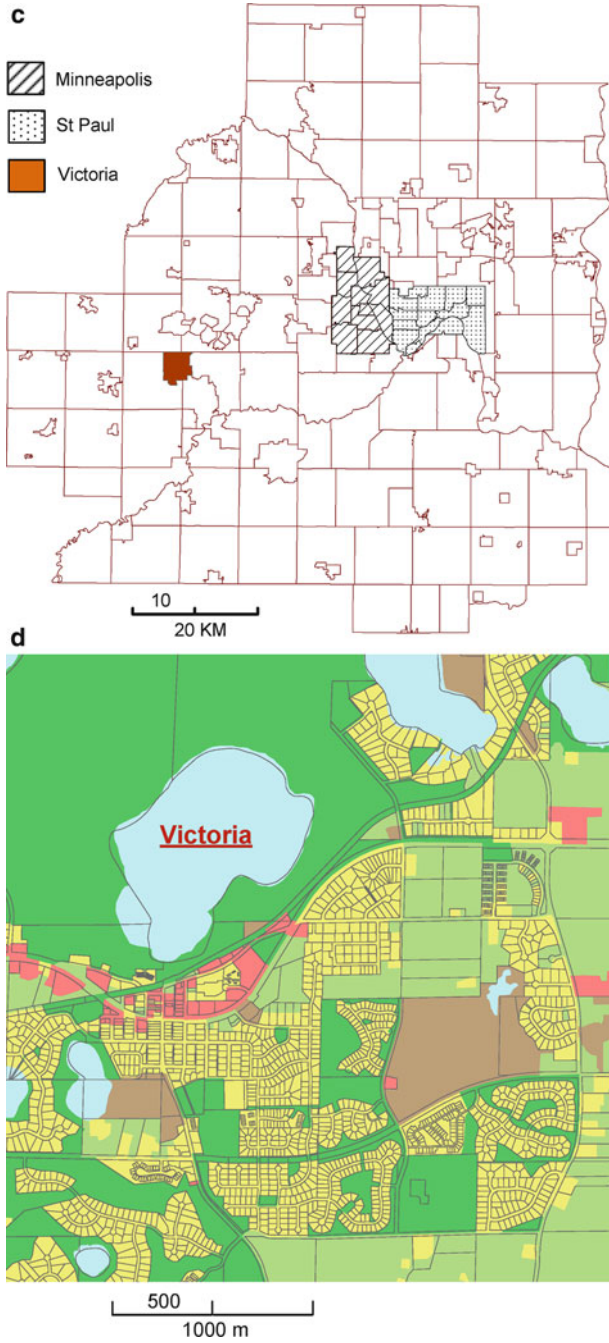


Fig. 6.3 (continued)

a random parcel within their neighborhood, but only one household can occupy a parcel. Agent decision making is defined by the conceptual model developed above, where households assess the probability of choosing a given vacancy (P_{ij}) via either distance-only or distance-plus-direction strategies. Households assess vacancies comprised of existing vacancies and recently developed houses, choosing vacancies per their move distance (given as a negative exponential distribution) and directional bias (per a von Mises distribution) as specified by the migration chain data.

6.3.2.3 Step 3: Simulate Migration

A key advantage of ABM is that they can be straightforward to run; for once agents are specified, and they are simply set in motion and dynamically interact with each other and the environment. Each model year, three processes occur. First, institutions apply policy rules on which areas can and cannot be developed. Second, developers build houses in developable locations that are added to the vacancy list. Third, households migrate, per direction and distance-and-direction rules, to new parcels and place their old houses on the vacancy list. Based on actual moves given by the parcel data, the estimated λ in Eq. 6.1 for distance in the TCMA is 0.160 for 2005–2007 and the estimated κ is 0.085 for Eq. 6.2 for direction (Fig. 6.4).

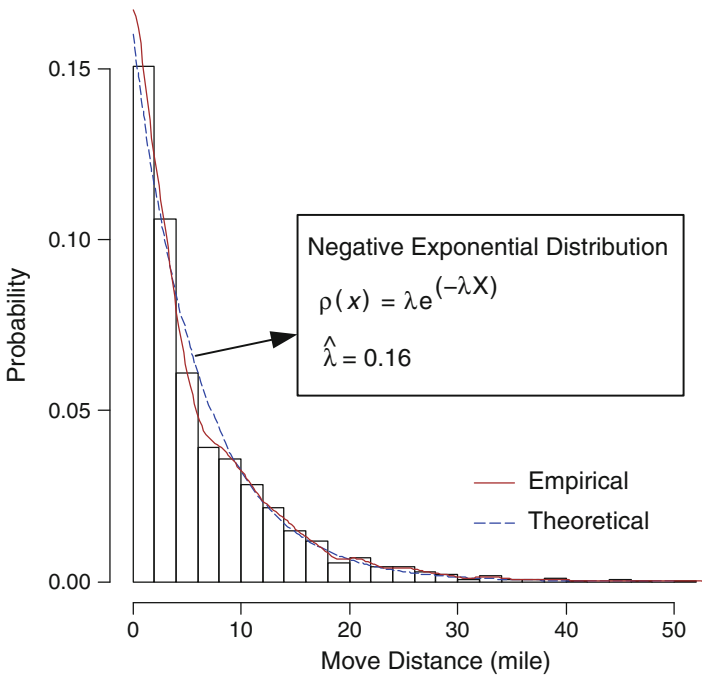


Fig. 6.4 Empirical distribution of move distance-and-direction

6.3.2.4 Step 4: Model Validation

Model validation involves measuring how well the model duplicates real-world phenomena. Model validation in the absolute or predictive sense is theoretically infeasible as no single model can reproduce every aspect of a complex open system (Oreskes 1998). That said, statistical measures can provide a useful benchmark for assessing how well different complex model configurations perform (Windrum et al. 2007; Manson 2007). Validation requires comparing model results to empirical data, to which end we used three different metrics: inner-migration rates, Syrjala tests, and minimum spanning trees. As model migration rates are calculated from actual relocation data, the number of modeled movers equals the actual number of migrations and values of λ in Eq. 6.1 and κ for Eq. 6.2 for direction. The main difference is therefore the spatial distribution of these migrant households. The three approaches employed to assess spatial fit are well suited to the problem at hand given that point pattern methods vary in their sensitivity and accuracy, as determined by their capacity to discriminate between point patterns, remain stable over different samples, and deal a range of underlying distributions (Wallet and Dussert 1998).

The three model validation approaches offer specific advantages while complementing one another. First, inner-migration rates compare the percentage of households that move within housing submarkets (i.e., those that stay within a given area or neighborhood). We use a multiscalar model specification across several submarket specifications to develop a strong measure of comparison between modeled and actual migration. Second, Syrjala tests compare the spatial distribution patterns of simulated and actual destinations of migrant households, offering the advantage over many standard point pattern analyses in assessing not just locations but also quantities across several scales of aggregation via a modified procedure that apportioned simulated and actual destination points. Third, use of minimum spanning trees (MST) offers an optimized nearest-neighbor distance analysis that, instead focusing on local nearest neighbors, describes the shortest, noncircular path connecting all points. Each of these three approaches offers distinct advantages as well as overlaps in validating how well simulated and actual migration match.

6.4 Results

In order to compare the simulated results with the actual distribution of migration destinations, we employ inner-migration rate comparison alongside Syrjala tests and MST to compare the similarity between the spatial distributions of actual and simulated migration destinations. Both distance and distance-and-direction yield realistic moves across scales of aggregation. Inner-migration rates at various scales indicate that the model recreates realistic aggregate spatial patterns of intraurban migration. Inner-migration rates measure the percentage of migrants who remain in the originating spatial unit and indicate the extent to which simulated moves match real relationships among vacant housing supply, move distance distribution, and

residential locations. In calculating inner-migration rates, it is necessary to correct for the fact that inner-migration rates are defined by arbitrary spatial units (Turner et al. 1989), in that small-area data can be combined at different resolutions. We measured the inner-migration rates across a series of 29 regular grids ranging from a coarse one 2×2 of grid cells to a fine-scaled 3×3 grid for the entire TCMA. At finer scales, up to a third of the grid cells fall outside of the seven-county region given its irregular boundary and are not included in the count because they would inflate the number of seemingly correct moves.

The simulated inner-migration rates mirror actual rates given by the parcel data, which indicates that the model captures key relationships between vacant housing opportunities, move distance distribution, and land-use patterns (Fig. 6.5a). Both decision-making strategies—distance-only and distance-plus-direction—produce inner-migration rates that are close to the actual values. Distance-and-direction outperforms just distance, as illustrated by the total root mean squared errors, which compares how well the simulation does against actual moves measured by inner-migration rates (Fig. 6.5b).

Syrjala tests offer an advantage over inner-migration rates in that they demonstrate how well a simulated distribution resembles an actual one (Syrjala 1996). The Syrjala test compares the values of two sets of samples or, in the case of intraurban migration, destinations tessellated onto a regular grid. The test produces two measures, a Syrjala statistic and a p value. The Syrjala statistic measures the differences between the cumulative distribution functions of the two samples. The smaller the statistic, the closer the two sample distributions, while the p value indicates the probability that the two samples are from the same population and spatial distribution. There is no simple analytical solution for p ; instead, it is calculated through sample permutation and denotes the percentage of randomized permutations that have bigger Syrjala statistic than the sample. If p is 0.03, for example, only 3 % of these random distributions are more similar to the distribution of one sample A than the other sample B, implying the spatial distribution of A is statistically different from that of B.

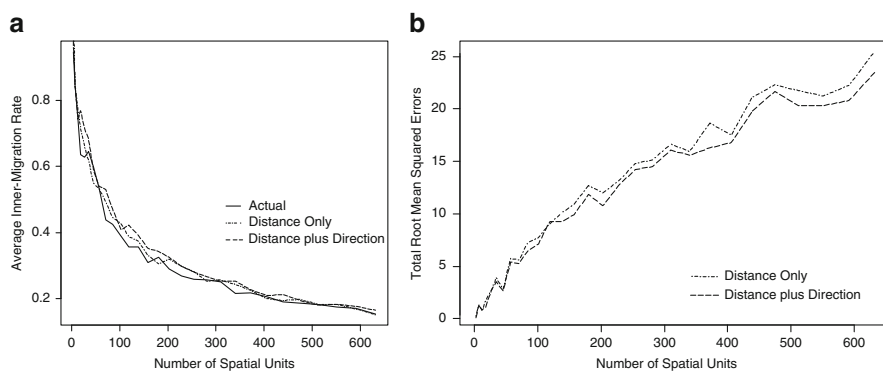


Fig. 6.5 Actual and simulated inner-migration rates (a) and total RMSE of simulated inner-migration rates against actual rates (b)

The Syrjala test of intraurban migration in the Twin Cities sheds light on complex patterns. Key measures are the percentage of subdivisions (not the number of households) that are not statistically different from the actual distribution (H_1), the mean Syrjala statistic \bar{S} , and the mean p value $p(\bar{S})$. First, both decision-making models score well on H_1 , where 71 % of subdivisions for the distance-only strategy and 69 % for the distance-plus-direction strategy match reality. Second, the distance-only strategy fares slightly better than distance-plus-direction strategy in recreating real migration patterns given lower average Syrjala statistics (0.696 vs. 0.771) and higher average p value (0.212 vs. 0.182). Overall, these two strategies are similarly successful in how they replicate real-world migration destinations.

The minimum spanning tree (MST) method focuses more on the relative position among intraurban migration destinations than the other two methods. Besides providing trees for visual inspection, the approach generates simple mean path length \bar{d} and variance $\sigma(d)$, where a short path length indicates that points are close to each other and a small variance means the points are evenly distributed. An MST is network structure that connects all nodes with a minimum total distance (Zahn 1971; West 2001). MSTs treat individual locations as nodes of a network in which each is connected to neighboring locations, which preserves information about the adjacency of nodes (Fig. 6.6). Importantly, an MST minimizes the length of the path connecting location while guaranteeing that every location linked to another one (Guo 2008). This approach preserves both absolute and topological spatial characteristics in a way that heightens sensitivity and accuracy (Wallet and Dussert 1998), as well as offering the benefit of identifying spatial hierarchies of migration.

Two specific examples illustrate how MST analysis compares the spatial distribution of simulated and actual migration destinations. In the exurban city of Norwood, migration extends toward the Minneapolis downtown (Fig. 6.7), which mirrors the simulated results. However, both decision-making models also produce two extra spurs on the MST that trend south and north, which is not consistent with the true situation.

For the inner-ring suburban city of Robbinsdale, simulated results have a more concentrated pattern than the real situation, implying that the average path length of simulated move destinations is shorter than the reality (Fig. 6.8).

A comprehensive comparison using MST features provides insights into the predictive powers of the two different decision-making strategies (Table 6.1). In terms of the mean shortest path length \bar{d} , the distance-only strategy produces the smallest minimum root of mean squared errors (RMSE) compared to actual migration. Both methods generate a smaller average path length than the real migration data, which means more compact patterning of moves. The lower value of the direction-plus-direction method compared with the distance-only method is expected because the directional bias compresses the migration destinations into a smaller region. The significantly shorter average path length of the distance-only method, together with the lower variance, implies that the distance-based methods tend to generate a more compact pattern than found in reality. In other words, they will underestimate urban growth and sprawl.

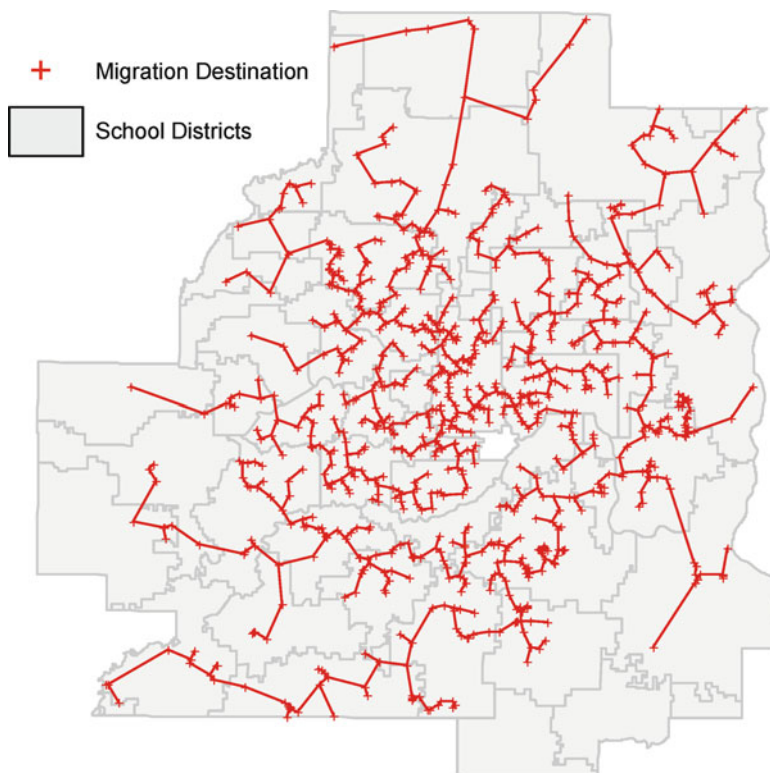


Fig. 6.6 Minimum spanning tree connecting intraurban migration destinations. Note: MST is scale independent. School districts serve as background

6.5 Discussion

Agent-based modeling of intraurban migration illustrates the importance of interaction between vacancy distribution and housing search, particularly in how complex intraurban migration patterns exhibited in the aggregate can arise from simple behavioral rules. The key finding of this work is that while it is a safe assumption that people take into account a range of personal, social, and environmental factors when making momentous housing decisions, distance-and-direction handily captures key facets of intraurban migration.

The prime import of this work is that it demonstrates a straightforward means for modeling the housing search process. The pure distance-based decision-making strategy, when applied to appropriately specified housing vacancies, can generate spatially realistic aggregate migration patterns. The addition of migration direction improves the fit somewhat at the cost of introducing greater complexity, given that it appears to capture the small but significant effect of directional bias even when using just a single downtown center instead of actual working places as the source

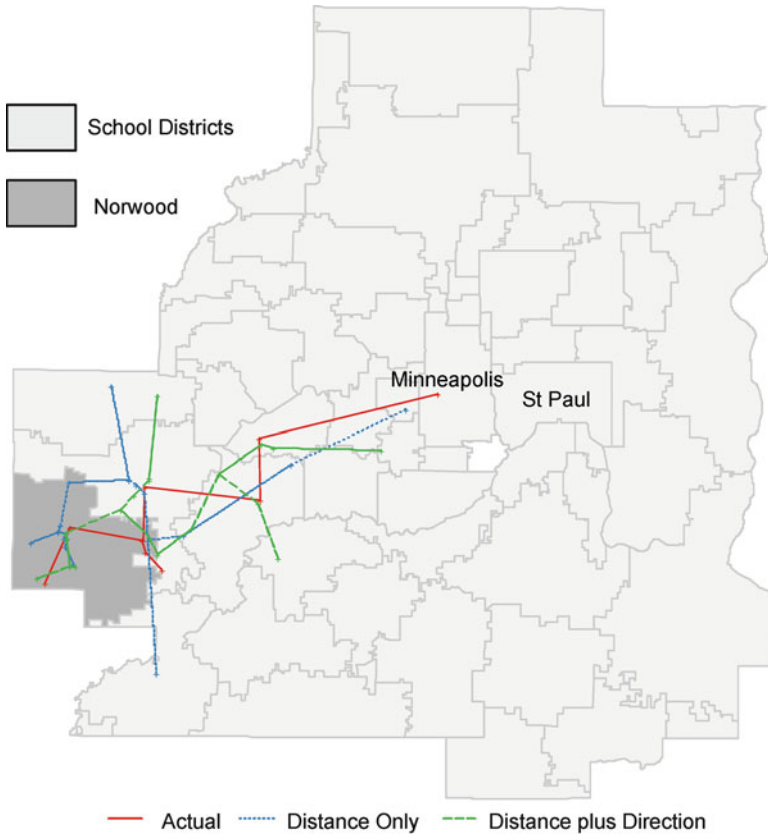


Fig. 6.7 Intraurban migration from Norwood

of this bias. Even then, underestimation of sprawl in suburban and exurban locales points to a prolonged housing search, which implies that people who live in areas with low population density tend to move less frequently and longer distances (see also Van der Vlist et al. 2002).

This work provides a basis for more complicated, utility-comparison-based migration models of housing search strategies. The modified intervening opportunity theory as instantiated in an agent-based model and empirically calibrated with migration chains captures fundamental features of the migration process and can complement deeper investigation of specific factors and locales. Alone, they can help capture key migration dynamics in the absence of many kinds of information usually needed to understand migration, creating a simple and powerful perspective on migration and urbanization. When combined with other data, they provide part of the foundation of a broader and deeper examination of intraurban migration. Directions for future research include more complicated spatial and social landscapes, such as a multi-nodal preference landscape or myriad public and

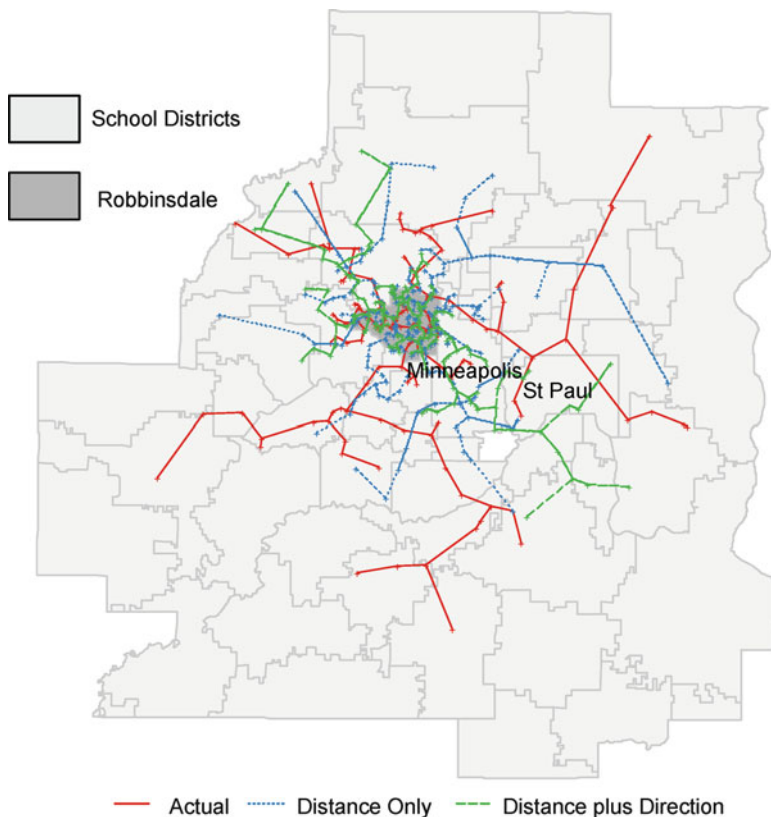


Fig. 6.8 Intraurban migration from Robbinsdale

Table 6.1 Comparison of migration strategy using MST path length distribution

Decision-making strategy	RMSE (m)	\bar{d}	$\sigma(d)$
Actual		4,962.08	4,626.66
Distance-only	3,908.477	3,655.031	3,042.45
Distance-plus-direction	5,051.047	3,042.447	2,713.88

private incentives related to housing, and consideration of how these landscapes interact with personal and household attributes.

More broadly, this work addresses a key methodological challenge for many urban modeling approaches, and especially for ABM, resisting the temptation to make models complicated. With data becoming more plentiful and methods growing increasingly sophisticated, models run the risk of committing what Lee (1973) termed the key “sins” of urban models, namely, being hyper-comprehensive and complicated at the cost of parsimony and generalizability (Lee 1973; Klosterman 1994). ABMs are at particular risk because their core strength is demonstrating

how complexity arises from actions and interactions of simple agents. There is a fundamental tension between the desire to create realistic models by incorporating many urban processes and the desire to explain how features of a city emerge from the simple interactions among entities such as households and properties (Clarke 2004; Brown et al. 2008). This tension gives rise to the need for simple, empirically based agent-based models of migration that can complement the host of more complicated models. Overall, this approach is deliberately straightforward in that it does not examine the characteristics of movers or the broader organization of housing—the primary foci of migration research—but instead centers on combining long-standing geographical findings to provide a straightforward sociospatial conceptualization of the intraurban migration process. Overall, this work complements existing approaches while breaking new ground in understanding how individual behavior scales up to the urban region.

This model also gives insight into how complexity emerges from simplicity by examining how specific housing opportunities and individual housing search behavior influence the aggregate pattern of intraurban migration. By combining intervening opportunities theory with behavioral evidence on the spatial characteristics of intraurban migration in an agent-based model, we can explore the extent to which real-world migration patterns can result from simple behavioral rules of household search in the context of housing opportunities. When households live in an area with fewer housing opportunities, for example, they are less likely to find a vacant house that meets their needs and thus require more iterations (i.e., more time) to accomplish their housing search. Importantly, while there are many different conceptual frameworks seek to explain migration, and while this diversity signifies healthy inquiry, it highlights the need for simple models of individual actions coupled to broader, generalizable (and admittedly simple) models of urban processes (Batty 2008, 2012). Methodological challenges abound, as evidenced by both the large array of statistical and simulation approaches used in migration analysis and the extent to which they are increasingly combined in hybrid models. Many of these theoretical and methodological issues have at their heart the need for better data, particularly on specific individuals and households who collectively drive intraurban migration. Taken together, these challenges indicate a pressing need for hybrid statistical and simulation models based on data on specific individuals to develop stronger conceptual frameworks of how individual actions give rise to the aggregate patterns and processes of intraurban migration.

Overall, while agent-based modeling can help explain complex systems by integrating many possible interacting components, it is also a valuable way to explore how straightforward behavioral rules of individuals can lead to processes and patterns of complexity. By examining, incorporating, and validating spatial behavioral theories, the modified intervening opportunities model offered here can serve as a sociospatial foundation for more comprehensive urban models as well as contribute to ongoing research on developing and validating theories of human behavior in urbanization.

Acknowledgments This work is supported in part by the National Aeronautics and Space Administration New Investigator Program in Earth-Sun System Science (NNX06AE85G), the National Institutes of Health supported Minnesota Population Center (R24 HD041023), and the Resident Fellowship program of the Institute on the Environment. The authors gratefully acknowledge the assistance of the editor and anonymous reviewers. Responsibility for the opinions expressed herein is solely that of the authors.

References

- Adams JS (1969) Directional bias in intra-urban migration. *Econ Geogr* 45(4):302–323
- Adams JS (1984) The meaning of housing in America. *Ann Assoc Am Geogr* 74(4): 515–526
- An L (2012) Modeling human decisions in coupled human and natural systems: review of agent-based models. *Ecol Model* 229:25–36
- Batty M (2008) The size, scale, and shape of cities. *Science* 319(5864):769–771
- Batty M (2012) Building a science of cities. *Cities* 29:S9–S16
- Benenson I, Torrens PM (2004) *Geosimulation: automata-based modeling of urban phenomena*. Wiley, Chichester
- Berger T, Schreinemachers P (2006) Creating agents and landscapes for multiagent systems from random samples. *Ecol Soc* 11(2):54–71
- Birkin M, Wu B (2012) A review of microsimulation and hybrid agent-based approaches. In: Crooks AT, See LM, Batty M, Heppenstall AJ (eds) *Agent-based models of geographical systems*. Springer, Dordrecht, pp 51–68
- Brown LA, Chung S-Y (2008) Market-led pluralism: rethinking our understanding of racial/ethnic spatial patterning in U.S. cities. *Ann Assoc Am Geogr* 98(1):180–212
- Brown LA, Longbrake DB (1970) Migration flows in intraurban space: place utility considerations. *Ann Assoc Am Geogr* 60(2):368–384
- Brown LA, Moore EG (1970) The intra-urban migration process: a perspective. *Geogr Ann B* 51(1): 1–13
- Brown DG, Robinson DT, An L, Nassauer JI, Zellner M, Rand W, Riolo R, Page SE, Low B, Wang Z (2008) Exurbia from the bottom-up: confronting empirical challenges to characterizing a complex system. *Geoforum* 39(2):805–818
- Choldin HM (1973) Kinship networks in the migration process. *Int Migr Rev* 10:163–175
- Clark WAV (1976) Migration in Milwaukee. *Econ Geogr* 52(1):48–60
- Clark WAV (1982) *Modeling housing market search*. St. Martin's Press, New York
- Clark WAV (1986) *Human migration*. Sage, Beverly Hills
- Clark WAV (2008) Geography, space, and science: perspectives from studies of migration and geographical sorting. *Geogr Anal* 40(3):258–275
- Clark WAV, Burt JE (1980) The impact of workplace on residential relocation. *Ann Assoc Am Geogr* 70(1):59–67
- Clark WAV, Flowerdew R (1982) A review of search models and their application to search in the housing market. In: Clark WAV (ed) *Modeling housing market search*. St. Martin's Press, New York, pp 4–29
- Clark WAV, Withers SD (1999) Changing jobs and changing houses: mobility outcomes of employment transitions. *J Reg Sci* 39:653–673
- Clark WAV, Huang Y, Withers S (2003) Does commuting distance matter? Commuting tolerance and residential change. *Reg Sci Urban Econ* 33(2):199–221
- Clarke KC (2004) The limits of simplicity: toward geocomputational honesty in urban modeling. In: Atkinson P, Foody G, Darby S, Wu F (eds) *GeoDynamics*. CRC Press, Boca Raton, pp 213–232

- Cochrane RA (1975) A possible economic basis for the gravity model. *J Transp Econ Policy* 9(1):34–49
- Cooke TJ (2008) Migration in a family way. *Popul Space Place* 14(4):255–265
- Crooks A, Castle C, Batty M (2008) Key challenges in agent-based modelling for geo-spatial simulation. *Comput Environ Urban Syst* 32(6):417–430
- De Jong GF, Roempke Graefe D (2008) Family life course transitions and the economic consequences of internal migration. *Popul Space Place* 14(4):267–282
- Diappi L, Bolchi P (2008) Smith's rent gap theory and local real estate dynamics: a multi-agent model. *Comput Environ Urban Syst* 32(1):6–18
- Dorigo G, Tobler W (1983) Push-pull migration laws. *Ann Assoc Am Geogr* 73(1):1–17
- Erlander S (2010) Behavioral foundations of spatial interaction models. In: Erlander SB (ed) *Cost-minimizing choice behavior in transportation planning*, *Advances in spatial science*. Springer, Berlin, pp 15–30. doi:[10.1007/978-3-642-11911-8_3](https://doi.org/10.1007/978-3-642-11911-8_3)
- Fernandez LE, Brown DG, Marans RW, Nassauer JI (2005) Characterizing location preferences in an exurban population: implications for agent-based modeling. *Environ Plan B: Plan Des* 32(6):799–820
- Fossett M (2006) Ethnic preferences, social distance dynamics, and residential segregation: theoretical explorations using simulation analysis. *J Math Soc* 30(3):185–273
- Fotheringham AS (1983) A new set of spatial-interaction models: the theory of competing destinations. *Environ Plan A* 15(1):15–36
- Gaile G, Burt J (1980) *Directional statistics. Concepts and techniques in modern geography*, vol 25. University of East Anglia, Norwich
- Geist C, McManus PA (2008) Geographical mobility over the life course: motivations and implications. *Popul Space Place* 14(4):283–303
- Gimblett HR (2002) *Integrating geographic information systems and agent-based modeling techniques for simulation of social and ecological processes*. Santa Fe Institute Studies on the Sciences of Complexity. Oxford University Press, New York
- Guldmann JM (1999) Competing destinations and intervening opportunities interaction models of inter-city telecommunication flows. *Pap Reg Sci* 78(2):179–194
- Guo D (2008) Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). *Int J Geogr Inf Sci* 22(7):801–823
- Haase D, Schwarz N (2009) Simulation models on human–nature interactions in urban landscapes: a review including spatial economics, system dynamics, cellular automata, and agent-based approaches. *Living Rev Landsc Res* 3(2)
- Irwin E, Jayaprakash C, Munroe D (2009) Towards a comprehensive framework for modeling urban spatial dynamics. *Landsc Ecol* 24(9):1223–1236
- Jackson J, Forest B, Sengupta R (2008) Agent-based simulation of urban residential dynamics and land rent change in a gentrifying area of Boston. *Trans GIS* 12(4):475–491
- Jayet H (1990) Spatial search processes and spatial interaction: 1. Sequential search, intervening opportunities, and spatial search equilibrium. *Environ Plan A* 22(5):583–599
- Jones C, Leishman C, Watkins C (2004) Intra-urban migration and housing submarkets: theory and evidence. *Hous Stud* 19(2):269–283
- Kennedy WG (2012) Modelling human behaviour in agent-based models. In: Crooks AT, See LM, Batty M, Heppenstall AJ (eds) *Agent-based models of geographical systems*. Springer, Dordrecht, pp 167–179
- Klosterman RE (1994) Large-scale urban models retrospect and prospect. *J Am Plan Assoc* 60(1):3–6
- Krzanowski R, Raper J (2001) *Spatial evolutionary modeling. Spatial information systems*. Oxford University Press, Oxford
- Kwasnicki W (1999) *Evolutionary economics and simulation*. In: Brenner T (ed) *Computational techniques for modelling learning in economics*. Kluwer Academic Publishers, Boston, pp 3–44
- Lee DB Jr (1973) Requiem for large scale urban models. *J Am Inst Plan* 39(3):163–178

- Loibl W, Toetzer T (2003) Modeling growth and densification processes in suburban regions – simulation of landscape transition with spatial agents. *Environ Model Softw* 18(6):553–563
- Macy MW, Willer R (2002) From factors to actors: computational sociology and agent-based modeling. *Annu Rev Sociol* 28:143–166
- Manson SM (2007) Challenges in evaluating models of geographic complexity. *Environ Plan B Plann Des* 34(2):245–260
- Manson SM, Evans T (2007) Agent-based modeling of deforestation in southern Yucatán, Mexico, and reforestation in the Midwest United States. *Proc Natl Acad Sci* 104(52):20678–20683
- Manson SM, Sander HA, Ghosh D, Oakes JM, Orfield MW Jr, Craig WJ, Luce TF Jr, Myott E, Sun S (2009) Parcel data for research and policy. *Geogr Compass* 3(2):698–726
- Matthews R, Gilbert N, Roach A, Polhill J, Gotts N (2007) Agent-based land-use models: a review of applications. *Landsc Ecol* 22(10):1447–1459
- Miller EJ, Douglas Hunt J, Abraham JE, Salvini PA (2004) Microsimulating urban systems. *Comput Environ Urban Syst* 28(1–2):9–44
- Mulder CH (2007) The family context and residential choice: a challenge for new research. *Popul Space Place* 13(4):265–278
- O’Sullivan D (2002) Toward micro-scale spatial modeling of gentrification. *J Geogr Syst* 4(3): 251–274
- O’Sullivan D (2008) Geographical information science: agent-based models. *Prog Hum Geogr* 32(4):541–550
- O’Sullivan D, Millington J, Perry G, Wainwright J (2012) Agent-based models – because they’re worth it? In: Crooks AT, See LM, Batty M, Heppenstall AJ (eds) *Agent-based models of geographical systems*. Springer, Dordrecht, pp 109–123
- Oreskes N (1998) Evaluation (not validation) of quantitative models. *Environ Health Perspect* 106(Suppl 6):1453
- Parker DC, Manson SM, Janssen M, Hoffmann MJ, Deadman PJ (2003a) Multi-agent systems for the simulation of land use and land cover change: a review. *Ann Assoc Am Geogr* 93(2): 316–340
- Parker DC, Manson SM, Janssen MA, Hoffmann MJ, Deadman P (2003b) Multi-agent systems for the simulation of land-use and land-cover change: a review. *Ann Assoc Am Geogr* 93(2): 314–337
- Pellegrini PA, Fotheringham AS (2002) Modelling spatial choice: a review and synthesis in a migration context. *Prog Hum Geogr* 26(4):487–510
- Quigley JM, Weinberg DH (1977) Intra-urban residential mobility: a review and synthesis. *Int Reg Sci Rev* 2(1):41
- Roseman CC (1971) Migration as a spatial and temporal process. *Ann Assoc Am Geogr* 61(3): 589–598
- Rouwendal J, Rietveld P (1994) Changes in commuting distances of Dutch households. *Urban Stud* 31(9):1545–1557
- Ruiter ER (1967) Toward a better understanding of the intervening opportunities model. *Transp Res* 1(1):47–56
- Salvini P, Miller EJ (2005) ILUTE: an operational prototype of a comprehensive microsimulation model of urban systems. *Netw Spat Econ* 5(2):217–234
- Simmons JW (1968) Changing residence in the city: a review of intraurban mobility. *Geogr Rev* 58(4):622–651
- Simpson L, Gavalas V, Finney N (2008) Population dynamics in ethnically diverse towns: the long-term implications of immigration. *Urban Stud* 45(1):163–183
- Smith TR, Clark WAV, Huff JO, Shapiro P (1979) A decision-making and search model for intraurban migration. *Geogr Anal* 11(1):1–22
- Stouffer SA (1940) Intervening opportunities: a theory relating mobility and distance. *Am Sociol Rev* 5(6):845–867
- Syrjala SE (1996) A statistical test for a difference between the spatial distributions of two populations. *Ecology* 77(1):75–80
- Torrens PM (2006) Simulating sprawl. *Ann Assoc Am Geogr* 96(2):248–275

- Torrens P (2012) Urban geosimulation. In: Heppenstall EAJ, Crooks AT, See LM, Batty M (eds) *Agent-based models of geographical systems*. Springer, Dordrecht, pp 435–450
- Torrens PM, Nara A (2007) Modeling gentrification dynamics: a hybrid approach. *Comput Environ Urban Syst* 31(3):337–361
- Turner MG, Costanza R, Sklar FH (1989) Methods to evaluate the performance of spatial simulation models. *Ecol Model* 48(1):1–18
- Van der Vlist A, Gorter C, Nijkamp P, Rietveld P (2002) Residential mobility and local housing-market differences. *Environ Plan A* 34(7):1147–1164
- van Ommeren J, Rietveld P, Nijkamp P (1997) Commuting: in search of jobs and residences. *J Urban Econ* 42(3402):421
- Waddell P (1993) Exogenous workplace choice in residential location models: is the assumption valid. *Geogr Anal* 25:65–82
- Wallet F, Dussert C (1998) Comparison of spatial point patterns and processes characterization methods. *Europhys Lett* 42:493–498
- West DB (2001) *Introduction to graph theory*, vol 2, 2nd edn. Prentice Hall, Upper Saddle River
- Windrum P, Fagiolo G, Moneta A (2007) Empirical validation of agent-based models: alternatives and prospects. *J Artif Soc Soc Simul* 10(2):8
- Wolpert J (1965) Behavioral aspects of the decision to migrate. *Pap Reg Sci* 15(1):159–169
- Zahn CT (1971) Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Trans Comput* 100(20):68–86
- Zaidi A, Rake K (2001) *Dynamic microsimulation models: a review and some lessons for SAGE*. SAGE discussion paper no. 2. Sage, London