Jul 1st, 12:00 AM

An Agent-based Model of Housing Search and Intraurban Migration in the Twin Cities of Minnesota

Shipeng Sun

Steven M. Manson

Sun, Shipeng and Manson, Steven M., "An Agent-based Model of Housing Search and Intraurban Migration in the Twin Cities of Minnesota" (2010). *International Congress on Environmental Modelling and Software*. 112.
https://scholarsarchive.byu.edu/iemssconference/2010/all/112

An Agent-based Model of Housing Search and Intraurban Migration in the Twin Cities of Minnesota

Shipeng Sun[a], Steven M. Manson[b]

a School of Planning, University of Waterloo, Canada Email: s27sun@uwaterloo.ca
b Department of Geography, University of Minnesota, USA Email: manson@umn.edu

Abstract: Intraurban migration defines many neighbourhood dynamics and consequently impacts land use patterns in the long term. Housing location decision-making is a complex process involving many features of the housing market that interact with the perceptions of home searchers. Although modellers have paid much attention to the prices and utilities of the environmental, ecological, and public services associated with housing, the housing search process is neglected in many agent-based land use models involving urban housing market. The challenge of incorporating housing search partially lies in the prohibitive cost of identifying, recording, and quantifying housing search activities at a large scale. This paper presents an agent-based model of intraurban migration featuring straightforward yet empirically accurate rules for housing search. Drawing on intervening opportunity and intraurban migration theories, this model is specified and calibrated using real-world housing vacancies and relocation origin-destination pairs extracted from parcel records available in the Twin Cities for 2005 to 2007. Multiple validation methods, including inner migration rates, Syrjala tests, and minimum spanning tree comparisons, show that the search rules based on housing vacancy distribution and negative exponential distance-decay probability can satisfactorily simulate the pattern of the housing search and locational choices made by homeowners in the Twin Cities of Minnesota.

Keywords: Intraurban migration, agent-based modelling, housing locational decisions

1. Introduction

Intraurban migration is a complex process involving many features of the housing market that interact with the perceptions of home searchers (Simmons 1968; Brown and Moore 1970; Roseman 1971; Dorigo and Tobler 1983; Clark 1986). Intraurban migration defines many neighbourhood dynamics, impacts land-use patterns, and consequently influences ecological systems in the long term as new residents in a neighbourhood may have different resources, different landscape and land-use preferences, and different land management practices. While migration behaviour shows clear and significant aggregate patterns such as suburbanization, it becomes less predictable when focusing on individual households, given the paucity of micro-scale public data available to researchers (Adams 1969; Clark 1976, 1986). Nevertheless, intraurban migration modelling at the individual level provides a way to actively explore the behavioural basis for household relocation decisions that complements other aggregate approaches like descriptive statistics and geovisualization.

Agent-based modelling of residential choice helps extend features of behavioural and economic models of intraurban migration, especially bridging simple micro behaviour and complex macro patterns. While agent-based modelling can embrace multiple elements in relocation decision-making and explore their nonlinear, complex interactions, it also provides a virtual laboratory to discover the simplicity of individual behaviour in terms of pinpointing rules of thumb that can depict much of the observed patterns and/or processes. Many agent-based urban models artfully explore and reveal the aggregate complexity of urban phenomena; however, it is also an ongoing challenge to integrate theoretical and empirical ingredients and in explaining the underlying simplicity of complexity (Miller et

al. 2004; Salvini and Miller 2005; Fossett 2006; Torrens 2006). There is a perennial tension between relying on empirical data and computational power to explain certain processes and patterns versus remaining theoretical and rarely handling real-world phenomena. Noticeably, with the advancement of computational techniques and the growing availability of fine-scale, spatially referenced socioeconomic and biophysical data, modellers are more likely than ever to commit the "sins" of large-scale, empirical models, such as hyper-comprehensiveness and complicatedness, (Lee 1973; Klosterman 1994).

To address these issues, this paper focuses on intraurban migration decisions made by homeowners in the Twin Cities metropolitan area (TCMA) of Minnesota, USA. By integrating migration theories and empirical evidence obtained from parcel data, we aim to discover features of the simplicity of individual decision-making that leads to complex migration patterns in aggregate. This paper draws on theories of intervening opportunity and relocation behaviour to find and verify such simple rules. We examine two factors in the housing search and relocation process that can generate the aggregate spatial patterns in the real world. One factor is the distribution of vacancies in the housing market; the other is the distance and direction between these vacancies and the mover's current dwelling. Our results can potentially provide a simple yet accurate way of describing housing search and migration processes. Unlike existing agent-based models of urbanization or land-use change that include relocation process based on utility comparison (e.g., Ettema et al. 2007), this paper specifically focuses on the spatial aspect of housing search and decision-making. It does not intend to address the motivation for relocation at the individual level. According to the two-stage model of intraurban migration, people make their decisions to move before starting the housing search process (Clark 1986). This paper therefore takes on the question of where people would move once they have decided to relocate within a metropolitan area and how well distance and directional bias alone can simulate the housing search process and the aggregate migration pattern.

The rest of the paper is organized as follows. The next section reviews locational decision-making theories on intraurban migration and proposes an empirically modified model for housing location decisions with two different strategies. Section three describes data and model specification, including the modelling procedure. Section four presents the model results and its validation. The paper concludes with discussion of our findings and their implications for general agent-based modelling.

2. Housing Search and Locational Decision Making

While intraurban migration has multiple components, its core is how households make their decisions about housing locations. Many theories of intraurban migration, especially in economics, focus on utilities and constraints; behavioural theories emphasize the importance of how homebuyers perceive and search the vacancies in the local and regional housing markets (Smith et al. 1979; Clark 1982; Clark and Flowerdew 1982; Dieleman 2001). Among others, particularly germane to migration are theories of intervening opportunities, distance-decay, and directional bias (Stouffer 1940; Adams 1969).

2.1 Theories on Housing Perception and Housing Search

As described by Wolpert (1965) and expanded on over the years, household migration can be seen as having three interrelated components — 1) social, demographic, economic, and environmental conditions that trigger household migration, 2) utilities of the current housing and expected utilities of other housing opportunities on the regional market, and 3) housing perceptions shaped by information collected through various communication paths. The model presented in this article concentrates on the third component.

Stouffer (1940) developed an intervening opportunity model by positing a mathematical relationship between housing opportunities and moving distances within a metropolitan area. With an assumption in which the quantity of vacant housing units is proportional to the distance from a household's current dwelling, Stouffer deduced that the number of households that move a certain distance has a logarithmic relationship with the housing opportunities located within that particular distance. This theory implies that people are unlikely to move further if they can find a vacancy near their current dwelling, which is

consistent with the phenomenon of short-distance domination in intraurban migration (Clark 1986). Stouffer verified the model with the 1930 intraurban migration data in the Cleveland metropolitan area at the census tract level, and it evolved into the influential gravity model and other spatial interaction models (Ruiter 1967; Cochrane 1975; Fotheringham 1983; Jayet 1990; Guldmann 1999). This work complements finding on more general patterns for moving distances; in particular, Quigley, Clark, and others formulated the exponential distribution of moving distances in intraurban migration (Quigley and Weinberg 1977; Clark and Burt 1980).

When people move, they consider not only distance but also directionality. Adams (1969) noted that people develop a sectoral bias extending from their dwelling to working places in their perception of a region. Using intraurban migration information extracted from city phone and address directories, Adams examined the migration patterns in Minneapolis and argued that the spatial search and residential locational choice of households are based on a limited mental map or image of the city. More importantly, the image is sector specific, namely, a narrow, wedge-shaped image with more focus on areas close to home, a finding validated by later research (Clark and Burt 1980; Clark, Huang, and Withers 2003).

2.2 An Empirically Modified Intervening Opportunity Model for the TCMA

This research adopts two housing search and relocation strategies for households that distil the essence of intraurban migration theories noted above — distance-only and distance-plus-direction. These two strategies are conceptually similar to the intervening opportunity model but are conditioned by the statistical distribution of moving distances and directions (Clark, Huang, and Withers 2003). In this sense, this paper introduces a new joint decision making model as illustrated by Figure 1.

In the model, two separate lists record all potential homebuyers and vacant houses on the regional housing market (Figure 1). In each model iteration, a homebuyer is randomly chosen to search the vacancy list. All houses in the vacant list have a random probability of being chosen by this particular homebuyer. The agent generates a random number, and the vacant house with a closest greater probability (calculated below) is chosen as the destination. When homebuyers move into vacant houses, they are removed from the buyer list and their previous house is added into the vacant house list.

With the distance-only strategy, each actor agent calculates the probability associated with vacant houses based on the distance between her current dwelling and a particular vacant house. More specifically, the probability follows a negative exponential distribution. Assume a homebuyer B_i now lives in H_i and the probability that she chooses house H_j would be $P_{ij} = \lambda e^{-\lambda d(H_i, H_j)}$, where λ is a parameter estimated from empirical move distance distribution in the TCMA and $d(H_i, H_j)$ is the distance between H_i and H_j.

With the distance-plus-direction strategy, not only is relocation distance considered but directional bias is also included in the calculation of probability P_{ij}. Although some studies show that long moves are more likely to occur along the home-downtown corridor, when the relocation process 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, Huang, and Withers 2003). In order to simulate the bi-modal distribution of move directions, the circular normal distribution, i.e., von Mises distribution is modelled as two normal distributions with zero and 180 degrees as mean values respectively. As a result, $P_{ij} = \lambda e^{-\lambda d(H_i, H_j)} \cdot P_\theta$, where P_θ is the probability that a homebuyer move in the direction of θ. 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 the standard normal distribution. When $\text{Sign}(\theta)$ is one, homebuyers move toward suburbs; when it is zero, they move toward downtown areas. The standard deviation σ_1 and σ_2 control the extent to which migrant households concentrate along home-downtown corridor when they move. When these deviation parameters are small, houses near the corridor have greater probabilities of being chosen by migrant households. When they are large, a greater number of houses have increasingly

similar odds of being chosen. When σ_1 is greater than σ_2 , households are more likely to move to suburbs; when smaller, households move downtown.
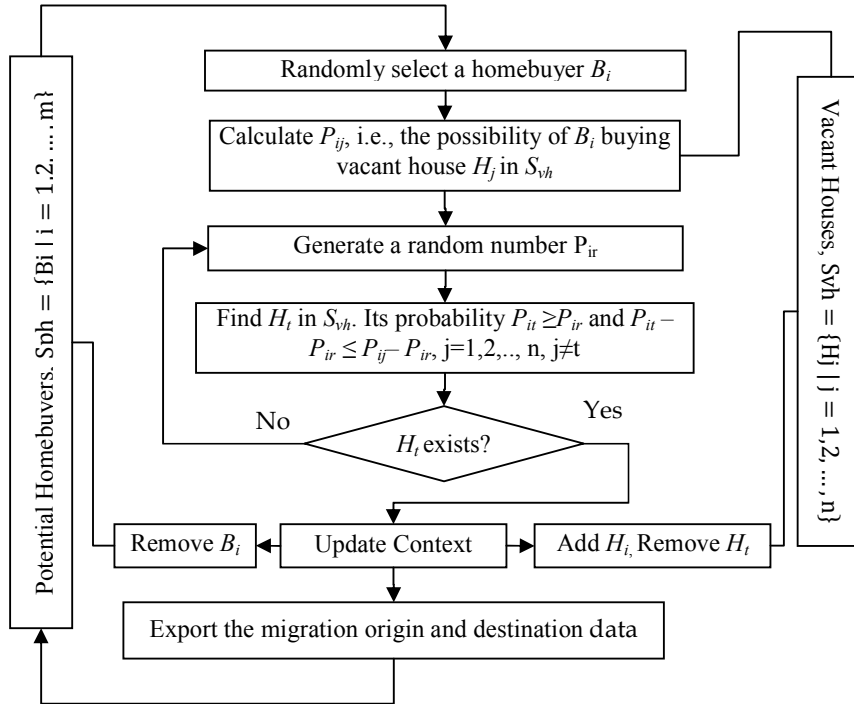

Figure 1 Intraurban Migration Decision-making Process

By combining the classic intervening opportunities model with behavioural evidence on the spatial characteristics of intraurban migration, the aggregate pattern of actor agents' relocations follows the expected statistical distribution of moving distance and direction. More importantly, as this model is implemented in an agent-based model, the actions of households are shaped by housing opportunities on the ground. When households are in an area with fewer housing opportunities, for example, they are less likely to find a vacant and require more iterations to finish their housing search, which is in line with findings that people who live in areas with low population density tend to move less frequently and longer distances (Van der Vlist et al. 2002). Overall, the agent-based model is able to examine how housing opportunities and housing search behaviours influence, and in some ways determine, the aggregate pattern of intraurban migration, leaving aside the decision of any given agent to move. This focus allows us to explore the extent to which real-world migration patterns can result from simple behavioural rules of households searching housing opportunities in the regional market.

3.   Data and Model Specification

The intraurban migration data used for theory development, model specification and validation are extracted from parcel data. These data describe the spatial location and characteristics of individual plots of land for tax purposes for the seven counties in the TCMA, and spans 2002 to 2009. Parcel data are valuable because they represent a near complete list of homeownership. Whenever parcel ownership changes, it is reasonable to infer that the previous owner likely moved out while the current owner moved in during the last year. We extracted all such changes and identified likely homeownerships changes by weeding out speculation, bank sales, and other transfers that did not involve owner-occupiers. Our data indicate that the revealed preference for housing during relocation has much randomness at the individual level due to the diversity of the homeowners and the size of housing stock in the TCMA. As such, the individual level preference has smaller impacts on the resultant spatial relocation pattern than housing search behaviour when

taken together. We therefore model homebuyers as agents with no socioeconomic features and no social interactions to examine how distance and directional bias alone produce the aggregate spatial pattern of intraurban migration. We fit the move distance and directional bias against the negative exponential and von Mises distribution using the statistical package, R. Parameters estimated from these distributions, including average move distance and directional bias, are adopted for model specification and calibration. For 2005-2006, we identified about 4,800 origin-destination pairs, and they are used for model validation.

This model is implemented with Human-Environment Land-Integrated Assessment (HELIA), an agent-based modelling software suite programmed in Java. The process of modelling TCMA intraurban migration in HELIA has four steps: 1) specifying a spatial configuration of agents and environmental data; 2) feeding these data into the model and populating it with agents; 3) assigning behavioural rules to actor agents and simulating intraurban migration in line with the specification above; and 4) verifying and validating the model against empirical data.

4.   Model Results

The model produces a set of modelled moves that can be compared against actual moves. As the number of modelled movers maps onto the actual number of migrants, the main difference, therefore, is the spatial distribution of these migrant households. In order to compare the simulated results with the actual distribution of migration destinations, it is necessary to employ several mathematical and statistical measures to help assess model fitness. While model validation in the absolute or predictive sense is theoretically infeasible for complex open systems such as urban system (Oreskes 1998), statistical measures can provide a useful benchmark for assessing how well the model performs. We used three approaches: inner-migration rate comparison and two point pattern analysis techniques, the Syrjala test and the minimum spanning tree (MST) method.

4.1   Inner-Migration Rates

Inner-migration rates at various scales are measured to ensure the models correctly incorporate the aggregate spatial patterns of intraurban migration. Inner-migration rates measure the percentages of migrants who remain in the same spatial unit. In calculating inner-migration rates, it is necessary to correct for the modifiable area unit problem given that inner-migration rates are defined by arbitrary spatial units (Turner, Costanza, and Sklar 1989). This problem is less of an issue in this study given that we are examining a large number of moves; but to offset the aggregation problem, we defined a series of regular grids to measure the inner-migration rates within them. More specifically, 29 grid systems are generated with 2×2 to 30×30 grids for the entire TCMA.
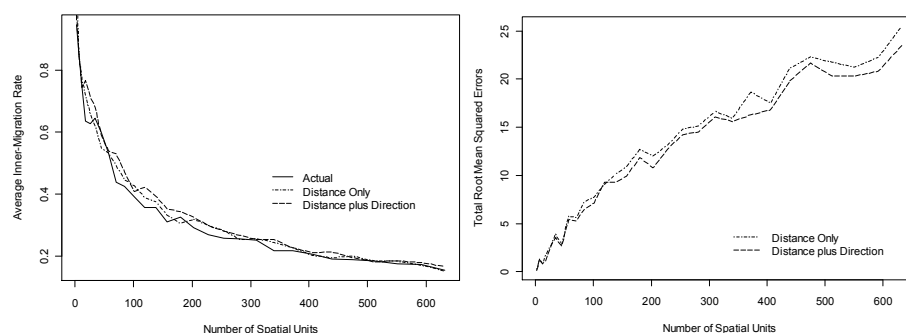

Figure 2 (a) Inner-migration Rates and (b) Total RMSE

The resultant simulated inner-migration rates are very close to the actual rates. This comparison suggests that the overall modelling framework correctly evaluates vacant housing opportunities, move distance distribution, and land use patterns (Figure 2a). Both residential location decision-making strategies—distance-only and distance-plus-direction—produce inner-migration rates that are close to the actual values. Distance-with-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) (Figure 2b).

4.2 Syrjala Test

While the inner-migration comparison can validate the agent-based model with two different housing locational decision-making strategies aggregately, the Syrjala test offers a means to tell how well the simulated distribution resemble the actual one over a series of arbitary sampling areas (Syrjala 1996). The Syrjala test essentially compares the values of two sets of samples at certain fixed locations via 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. The p value is a probability indicator of how likely these two samples are from the same population or how likely it is that they have the same spatial distribution.

The Syrjala comparison of real and modelled intraurban migration in Twin Cities shows complex patterns among four measures, namely the number of areas that have no statistically different distributions from the actual situation (H_1), the times for making the best prediction with a largest p value (N_{max}), the mean Syrjala statistics ($\bar{S}$), and the mean p value ($\overline{p(S)}$). First, both housing locational decision-making models generate satisfactory spatial patterns for migration destinations. In the 52 areas with a large enough sample size to calculate the Syrjala statistic, the null hypothesis that the modelled distribution is the same as the actual cannot be rejected in at least 71% of these areas for the distance-only strategy and 69% for the distance-plus-direction strategy (Table 1). Second, the distance-only decision-making strategy fares slightly better than distance-plus-direction strategy, of which the former has lower average Syrjala statistics (0.696 vs. 0.771) and higher average p value (0.212 vs. 0.182). The smaller Syrjala statistic of distance-only strategy means that migration destinations generated from this strategy is closer to their actual distribution than distance-plus-direction strategy. In addition, the distance-only method has a greater probability of having identically spatial distribution as actual situation because its p value is higher than distance-plus-direction. Overall, however, these two strategies are statistically similar to each other in terms of replicating real-world migration destinations.

Table 1 Model Comparison using Syrjala Test

Decision-making strategy	H_1	N_{max}	$\bar{S}$	$\overline{p(S)}$
Distance Only	37	6	0.696	0.212
Distance plus Direction	36	6	0.771	0.182

Notes: The number of areas with sufficient sample size is 52.

4.3 Minimum Spanning Tree Method

Minimum spanning tree (MST) methods focus on the relative position of intraurban migration destinations. By connecting all destinations using the shortest path, the minimum spanning tree can reveal the internal structure of these points. The simple mean path length ($\bar{d}$) and variance ($\sigma(d)$) describe the relative location of these points in the tree. Shorter average path length indicates that points are closer to each other and a smaller variance means the points are more evenly distributed.

A comprehensive comparison using MST features provides insights into the predictive powers of the two different decision-making strategies (Table 2). In terms of the mean shortest path length, $\bar{d}$, the distance-only decision-making 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 pattern of moves. The relative lower value of direction-plus-direction method compared with distance only method is expected because the directional bias actually compresses the migration destinations into a smaller region. The significantly shorter average path length of distance-only method, together with the lower variance, probably implies that the distance-based methods tend to generate a more compact pattern than the reality. In other words, they will underestimate real-world urban growth and sprawl.

Table 2 Comparison Using Path Length Distribution in MST

Decision-making Strategy	RMSE (m)	$\bar{d}$	$\sigma(d)$
Actual		4962.08	4626.66
Distance Only	3908.477	3655.031	3042.45
Distance plus Direction	5051.047	3042.447	2713.88

5. Conclusion and Discussion

Agent-based modelling of intraurban migration illustrates the importance of housing vacancy distributions and the role of distance and direction in developing housing perceptions. The pure distance-based decision-making strategy, when tied to housing vacancies, can generate aggregate migration patterns consistent with reality. In other words, complex intraurban migration patterns seen at the aggregate level can result from simple behavioural rules. The addition of migration direction improves the fit somewhat at the cost of greater complexity, given that it appears to capture a small yet significant effect of directional bias even when using just downtown centres instead of actual working places as the source of directional bias.

For many urban modelling applications that focus on aspects of urban morphology other than migration, it is possible to use the negative exponential distribution-controlled distance strategy as a proxy to real world behaviour. For those that do explicitly include household relocation, this empirical model also provides a theoretically derived and empirically specified prototype for the housing search process. Given that homebuyers cannot visit and evaluate all vacancies in the housing market, this work complements more complicated, utility-comparison based migration models that invoke dozens of variables and complicated decision making strategies. In this sense, the simpler approach developed here offers insight into the role of incomplete information and bounded rationality in relocation. More generally, while agent-based modelling can help understand and explain complex systems by integrating as many as possible interacting components, it is also valuable to explore the simple yet significant behavioural rules that influence such complexity. By examining and incorporating spatial behavioural theories, researchers and modellers can potentially contribute to building more comprehensive models as well as developing and validating new theories. For agent-based modelling practitioner, achieving balance between complexity and simplicity and between being theoretical and being empirical might better serve members in the broader research community by presenting more accessible results and conclusions.

REFERENCES

Adams, J. S. 1969. Directional Bias in Intra-Urban Migration. *Economic Geography* 45 (4):302-323.

Brown, L. A., and E. G. Moore. 1970. The Intra-Urban Migration Process: A Perspective. *Geografiska Annaler* B (25).

Clark, W. A. V. 1976. Migration in Milwaukee. *Economic Geography* 52 (1):48-60.

———. 1982. *Modeling Housing Market Search*. New York: St. Martin's Press.

———. 1986. *Human Migration*. Beverly Hills: Sage Publications.

Clark, W. A. V., and J. E. Burt. 1980. The Impact of Workplace on Residential Relocation. *Annals of the Association of American Geographers* 70 (1):59-67.

Clark, W. A. V., and R. Flowerdew. 1982. A Review of Search Models and Their Application to Search in the Housing Market. In *Modeling Housing Market Search*, edited by W. A. V. Clark. New York: St. Martin's Press.

Clark, W. A. V., Y. Huang, and S. Withers. 2003. Does Commuting Distance Matter? Commuting Tolerance and Residential Change. *Regional Science and Urban Economics* 33 (2):199-221.

Cochrane, R. A. 1975. A Possible Economic Basis for the Gravity Model. *Journal of Transport Economics and Policy* 9 (1):34-49.

Dieleman, F. M. 2001. Modelling Residential Mobility: A Review of Recent Trends in Research. *Journal of Housing and the Built Environment* 16 (3-4):249-265.

Dorigo, G., and W. Tobler. 1983. Push-Pull Migration Laws. *Annals of the Association of American Geographers* 73 (1):1-17.

Ettema, D., K. Jong, H. Timmermans, and A. Bakema. 2007. Puma : Multi-Agent Modelling of Urban Systems. In *Modelling Land-Use Change: Progress and Applications*, edited by E. Koomen, J. Stillwell, A. Bakema and H. J. Scholten: Springer Netherlands.

Fossett, M. 2006. Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations Using Simulation Analysis. *The Journal of Mathematical Sociology* 30 (3):185 - 273.

Fotheringham, A. S. 1983. A New Set of Spatial-Interaction Models: The Theory of Competing Destinations. *Environment and Planning A* 15 (1):15-36.

Guldmann, J. M. 1999. Competing Destinations and Intervening Opportunities Interaction Models of Inter-City Telecommunication Flows. *Papers in Regional Science* 78 (2):179-194.

Jayet, H. 1990. Spatial Search Processes and Spatial Interaction: 1. Sequential Search, Intervening Opportunities, and Spatial Search Equilibrium. *Environment and Planning A* 22 (5):583-599.

Klosterman, R. E. 1994. Large-Scale Urban Models Retrospect and Prospect. *Journal of the American Planning Association* 60 (1):3 - 6.

Lee, D. B., Jr. 1973. Requiem for Large Scale Urban Models. *Journal of the American Institute of Planners* 39 (3):163-178.

Miller, E. J., J. Douglas Hunt, J. E. Abraham, and P. A. Salvini. 2004. Microsimulating Urban Systems. *Computers, Environment and Urban Systems* 28 (1-2):9-44.

Oreskes, N. 1998. Evaluation (Not Validation) of Quantitative Models. *Environmental Health Perspectives* 106 (Suppl 6):1453.

Quigley, J. M., and D. H. Weinberg. 1977. Intra-Urban Residential Mobility: A Review and Synthesis. *International Regional Science Review* 2 (1):41-66.

Roseman, C. C. 1971. Migration as a Spatial and Temporal Process. *Annals of the Association of American Geographers* 61 (3):589-598.

Ruiter, E. R. 1967. Toward a Better Understanding of the Intervening Opportunities Model. *Transportation Research* 1 (1):47-56.

Salvini, P., and E. J. Miller. 2005. Ilute: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems. *Networks and Spatial Economics* 5 (2):217-234.

Simmons, J. W. 1968. Changing Residence in the City: A Review of Intraurban Mobility. *Geographical Review* 58 (4):622-651.

Smith, T. R., W. A. V. Clark, J. O. Huff, and P. Shapiro. 1979. A Decision-Making and Search Model for Intraurban Migration. *Geographical Analysis* 11 (1):1-22.

Stouffer, S. A. 1940. Intervening Opportunities: A Theory Relating Mobility and Distance. *American Sociological Review* 5 (6):845-867.

Syrjala, S. E. 1996. A Statistical Test for a Difference between the Spatial Distributions of Two Populations. *Ecology* 77 (1):75-80.

Torrens, P. M. 2006. Simulating Sprawl. *Annals of the Association of American Geographers* 96 (2):248-275.

Turner, M. G., R. Costanza, and F. H. Sklar. 1989. Methods to Evaluate the Performance of Spatial Simulation Models. *Ecological Modelling* 48 (1):1-18.

Van der Vlist, A., C. Gorter, P. Nijkamp, and P. Rietveld. 2002. Residential Mobility and Local Housing-Market Differences. *Environment and Planning A* 34 (7):1147-1164.

Wolpert, J. 1965. Behavioral Aspects of the Decision to Migrate. *Papers in Regional Science* 15 (1):159-169.