Examining Intraurban Migration in the Twin Cities Metropolitan Area using Parcel Data

Author's final proofread version of the Chapter 6 in "Big Data for Regional Science"

Shipeng Sun

Then:

Department of Environmental Studies
University of Illinois Springfield
PAC 314 One University Plaza
Springfield, IL 62703
Phone: (217) 206-7895
Email: sunsp.gis@gmail.com, ssun32@uis.edu

Now:

Department of Geography, Hunter College
Earth and Environmental Sciences, Graduate Center
The City University of New York
695 Park Ave
New York, NY 10065
Phone: (212) 396-6039
Email: sunsp.gis@gmail.com, shipeng.sun@hunter.cuny.edu

Citation:

Sun, S. (2017). Examining intraurban migration in the Twin Cities metropolitan area using parcel data. In L. A. Schintler & Z. Chen (Eds.), *Big Data for Regional Science* (pp. 64-74). London: Routledge. (APA Style)

Sun, Shipeng. 2017. "Examining intraurban migration in the Twin Cities metropolitan area using parcel data." In *Big Data for Regional Science*, edited by Laurie A Schintler and Zhenhua Chen, 64-74. London: Routledge. (Chicago Style)

Abstract

Intraurban migration, that is residential relocation within a metropolitan area, is an important aspect of regional and urban studies. However, the paucity of data on the migration choices of individuals at finer spatial scales remains a challenge. Using big land parcel data, a novel form of information on household intraurban migration was developed for the Twin Cities metropolitan area of Minnesota. With such information, the spatial patterns of migration as well as the changes of structural and neighborhood characteristics before and after migration were revealed and examined using spatial analysis and big data analytics.

Keywords: intraurban migration, parcel dataset, urban big data

Intraurban Migration

Migration is one of the three fundamental factors that drive population dynamics. Examining population migration at the international and national scales is increasingly important in a globalized economy (Castles, Miller, & Ammendola, 2005; Fan, 2002; Frey & Liaw, 2005). At the same time, research on intraurban migration, i.e., relocation within a metropolitan area, still retains its value in regional and urban studies for applications from regional housing market analysis, to residential structure change, to transportation planning, and to micro urban modeling (Brown & Moore, 1970; Clark, Deurloo, & Dieleman, 2006; De Jong & Roempke Graefe, 2008; Hipp & Boessen, 2016; Sun & Manson, 2012; Sun & Manson, 2015).

Research on intraurban migration, also called residential mobility, has a long history and is one classic topic in geography, sociology, economics, and other social sciences (Brown & Moore, 1970; Clark, 1976; Jones, Leishman, & Watkins, 2004). It mainly focused on three aspects: housing decision-making, constraints and context, and outcomes of intraurban migration. Housing decisions were mostly examined with classic economic or behavioral theories regarding how residents weight factors to make housing decisions (John S. Adams, 1969; Smith, Clark, Huff, & Shapiro, 1979). Constraints and context in intraurban migration research commonly pointed to housing affordability, housing information acquisition, and racial discrimination in housing search and lending practices (Galster, 1988). Geographical studies of intraurban migration outcomes concentrated on micro and macro aspects, i.e., household migration behavior and its aggregate impacts on residential structure (J. S. Adams, Caruso, Nordstrand, & Palm, 1973). Researchers had identified three aggregate spatial characteristics in the context of North American metropolitan areas: short-distance dominance (Haynes & Fotheringham, 1984), directional bias (John S. Adams, 1969), and suburban bias (Clark, 1976, 1986).

The paucity of data on the migration choices of individuals has always been a challenge in understanding intraurban migration. Big Data, the land parcel dataset in particular, provides an opportunity to derive large-scale individual intraurban migration information with precise locations (Kitchin, 2013; Rae & Singleton, 2015; Sun, 2009; Sun & Manson, 2012). Parcel data delineate individual property lots for tax purposes (Manson et al., 2009). Different local government agencies archive parcel data in databases with various variables, formats, quality standards, and metadata. They update the parcel dataset at different frequencies, often annually or quarterly. To extract useful and reliable intraurban migration information from parcel datasets, researchers must handle the variety and volume of parcel data. Due to input errors and frequent updates, the parcel dataset also has features of veracity and velocity. Overall, although parcel data have no comparison to the typical big data like those collected through mobile devices and sensors in real-time, they do have the core characteristics of big data and can substantially and more directly contribute to urban and housing studies (Kitchin, 2014; Wu, Wang, & Dai, 2016).

To illustrate the application of big parcel data, this chapter uses the Twin Cities Metropolitan Area (TCMA) of Minnesota as an example to examine intraurban migration. The TCMA parcel dataset was collected by the Metropolitan Council, which is the regional government of the TCMA. It had about one million records each year describing major attributes of land parcels related to property taxes. Using big data analytics and spatial analysis, this research extracted spatially accurate household relocation information and examined the spatial patterns and their relation to the associated socioeconomic and physio-environmental variables.

Using Big Parcel Data to Analyze Intraurban Migration

To obtain intraurban migration information, it is necessary to track household moves. There are four types of approaches: 1) general-survey data such as census microdata; 2) specific-survey data such as mail or telephone relocation surveys; 3) housing rental and sale transaction records; and 4) enumerative data such as city or telephone directories, utility bills, residence registry records, and parcel data (Clark, 1986; Komoto, 1994). While survey and transaction data offer much direct information about intraurban migration, existing enumerative data provide a less expensive yet more comprehensive alternative for migration analysis with near complete coverage of people, households, or houses in an area.

Parcel data describe the spatial location and characteristics of individual plots of land for tax purposes. The annual regional parcel dataset in the TCMA includes seven counties, namely Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. Each year, the Metropolitan Council collects parcel data for the previous year from these seven counties, unifies their format, merges them into a single dataset, and publishes them in GIS format. Although the Council made little effort to check data accuracy and consistency, parcel data generally have high quality because of its usage for taxes. The parcel dataset includes the contact information of owners or taxpayers, estimated property values, tax base, and the latest sale date and price. Since 2004, more structural attributes of the housing unit, like use type, lot size, garage, cooling, green space, and open space, had also been added, although their coverage and completeness varied across the seven counties. Over the years, the parcel dataset has included more complete information on parcel ID and owner or taxpayer names, the two most important attributes for migration information extraction. The percentage of records having valid parcel ID and names rose from 87% to 99% and the total housing stock had increased 9.9% in total for 2002 to 2007. The annual growth rate goes up from 1.63% to 2.90% for 2002 to 2006 but drops to

1.07% in 2007.

Because the annual parcel dataset contains homeownership for all parcels in the metropolitan area, comparing the owners of the same property over time gives information on who moved where and who occupied whose previous dwellings. It requires comparing the owners of a parcel across years, detecting valid owner changes, and matching the owner in previous and next year for moving in and moving out. The mechanism is implemented in a C# software package. For privacy and security reasons, the TCMA parcel data only contain names instead of identifiers like social security numbers. However, names are poor identifiers for people, families, or households and subject to duplication, multiform variations, and spelling errors. The influence of different homeowners having identical names was trivial because the joint probability that two owners had the same name and moved in the same year is extremely low. And relevant strategies were adopted in the software package to reduce these errors and ensure the quality of the extracted migration information (Sun, 2009). For instance, George Washington and G. Washington would be judged as the same person. George Washington, George & Martha Washington, and George Washington, Jr. would be the same household. Once the same names appeared in different properties across one or multiple years, moves were identified and the locations of the parcels and their socioeconomic and environmental characteristics were extracted using GIS. And then their spatial patterns and dynamics were examined using big data analytics and statistical analyses in R.

Application to the Twin Cities Metropolitan Area of Minnesota

Spatial patterns of intraurban migration

The spatial nature of migration can be studied through a migration vector pointing from origin to destination with links to work places (John S. Adams, 1969; Clark & Burt, 1980). The basic algebraic properties of such a vector are length (distance) and direction along the work-home corridor (Figure 6.1).

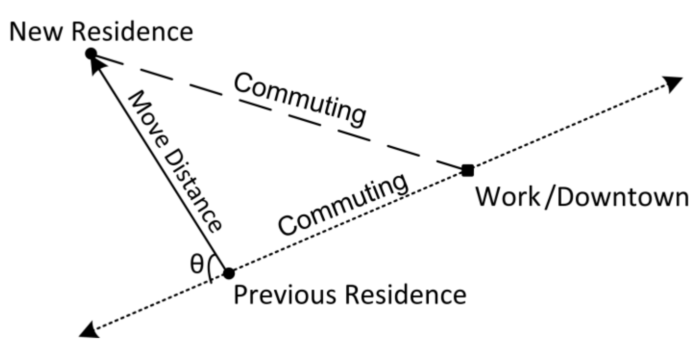

Figure 6.1 Intraurban Migration Vector.

Note: θ is the deviation from the work-home corridor.

Move distance and direction were commonly examined at aggregate scales, mostly at the census tracts level. The probability of moving a distance of d follows a negative exponential model, i.e., $p(d) = \lambda e^{-\lambda d}$, where λ is a parameter and its value is the reciprocal of the average move distance. The move directions follow a von Mises model, which is essentially a normal distribution for directional variables.

Comparing the empirical distance distribution and the one fitted using the exponential distribution shows some general patterns across space and time (Figure 6.2).
First of all, both short and long distance relocations are greater than
predicted by the exponential model while mid-range relocations are lower than expected. Specifically, the percentages of households that moved less than four or more than fifteen miles were higher than expected. The number of households who moved between four and fifteen miles, in contrast, was lower than expected. Using Roseman's classification of migration, the short-distance moves are partial displacement migrations that "involve displacement of only part of the everyday reciprocal movements of migrants" and try to minimize changes in other aspects (Roseman, 1971, p. 589). The long-distance moves are total displacement migrations with "complete spatial displacement of the daily/weekly reciprocal movement patterns" and are caused by job changes and migration from rural to urban areas (J. S. Adams et al., 1973; Mabogunje, 1970; Roseman, 1971, p. 589).

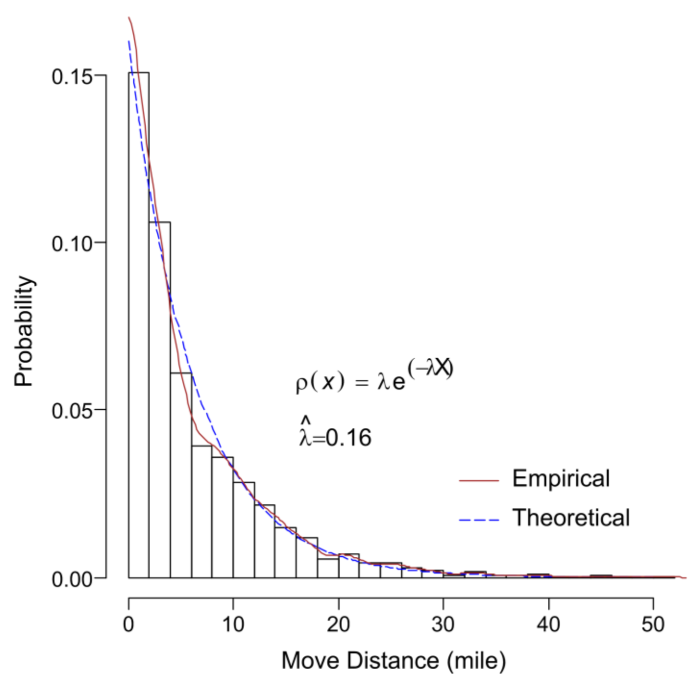

Figure 6.2 Intraurban Move Distance Distribution

Note: Move distances follow a negative exponential distribution with a parameter of λ.

The pattern also suggested that Weibull distribution would be a better model than the exponential for the move distances. Weibull distribution has one extra shape parameter and if the parameter is one Weibull becomes exponential. Statistical tests support the hypothesis that this parameter is significantly less than one (Table 6.1).
The fitted plots also clearly show that the Weibull distribution captures the short and long moves that are underestimated by the exponential model and can avoid overestimating the average move distances.

Table 6.1 Parameters Estimated from Intraurban Migration Distances

	Exponential Distribution			Weibull Distribution			
Year	$\hat{\lambda}$	$\bar{d}$	$\sigma(\lambda)$	$\hat{\alpha}$	$\sigma(\alpha)$	$\hat{\beta}$	$\sigma(\beta)$
2002-03	0.171	5.862	0.0023	0.86	0.0092	5.46	0.089
2003-04	0.158	6.329	0.0021	0.91	0.0098	6.05	0.095
2004-05	0.166	6.024	0.0024	0.81	0.0097	5.43	0.104
2005-06	0.161	6.231	0.0024	0.84	0.0103	5.74	0.109
2006-07	0.160	6.234	0.0033	0.87	0.0142	5.85	0.143

Notes: When $\alpha = 1$ and $\beta = 1/\lambda$, Weibull becomes exponential distribution.
The estimated $\hat{\alpha}$ is statistically different from 1.

Second, the parameter λ had a decreasing trend over time, which implied an increasing average move distance. This is easy to understand as people tend to move longer distances with the advancement of transportation technologies and the growth of metropolitan areas. The long term trend of moving distances is particularly evident. For example, the average move distance in Cleveland, OH was around 1.7 mile in 1920s, 1.9 miles in 1930s, and 2.2 miles in 1940s. This value jumped to 3.7 miles in Milwaukee, WI in 1960s and then to 6.28 in Seattle, WA in 1990s (Clark, 1976; Clark, Huang, & Withers, 2003). Although these values illustrate the general trend, they are not entirely comparable because the sizes of metropolitan areas are different. The migration information extracted from the annual parcel datasets, by contrast, provided a more accurate and more realistic delineation of the trend. From 2002 to 2007, both the Exponential and Weibull models suggest a rising trend in average move distances, although the year 2003-2004 had a high-than-expected value.

Besides move distances, the macro manifestation of intraurban migration directions is directional bias and the suburbanization trend (John S. Adams, 1969). Aggregately, people tend to move to suburban areas within a wedge-shaped area along their work-home corridors. Move directions can be modeled as a circular normal or a von Mises distribution, the counterpart of the normal distribution for directional data spanning 0 to 360 degrees. The density function of a von Mises distribution is $f(x|\mu,\kappa) = {e^{\kappa\cos(x-\mu)}}/{2\pi I_0(\kappa)}$, 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 as a normal distribution.

When it comes to empirically testing directional bias, existing studies arrived at no decisive conclusions (Clark, 1976; Clark & Burt, 1980; Clark et al., 2003; Whitelaw & Robinson, 1972). Using the downtown areas of Minneapolis and St. Paul as the approximate work places, the TCMA parcel dataset can help empirically test this theory. Although the κ is small, i.e., 0.05 ~ 0.10, they are significantly greater than zero at the 95% confidence level (Table 6.2) and the directional bias did seem to exist along the downtown-home corridor in the TCMA.

Table 6.2 Directional Bias Parameters Estimated from von Mises Distribution

year	$\hat{\mu}$	$\hat{\kappa}$	$[\mu$	$\mu]$	$[\kappa$	$\kappa]$
2002-03	0.3816	0.0650	-0.2085	0.9668	0.0260	0.1018
2003-04	0.4567	0.0575	-0.2908	1.1742	0.0178	0.0972
2004-05	0.3892	0.0720	-0.1446	0.9581	0.0312	0.1144
2005-06	2.9968	0.0877	2.4935	3.4872	0.0451	0.1316
2006-07	2.5926	0.0838	1.8347	3.3152	0.0259	0.1377

Notes: The confidence interval is at 95% level.

With directional bias, the tendency toward suburbs or downtown can be estimated as an indicator of suburbanization (Figure 6.3). The move directions
and distances are represented with windrose diagrams (Figure 6.3A). The orientation of each pedal in the diagram is indicative of the move direction, either towards downtown, towards suburbs, or with certain angles between the work-home corridor. The color of the pedal is indicative of the move distance. The size of the pedal is also proportional to the number of migrant households that moved within the specific distance and direction ranges. It is clear that short move distances dominate. When people move long, they tend to move towards the urban center instead of suburbs.

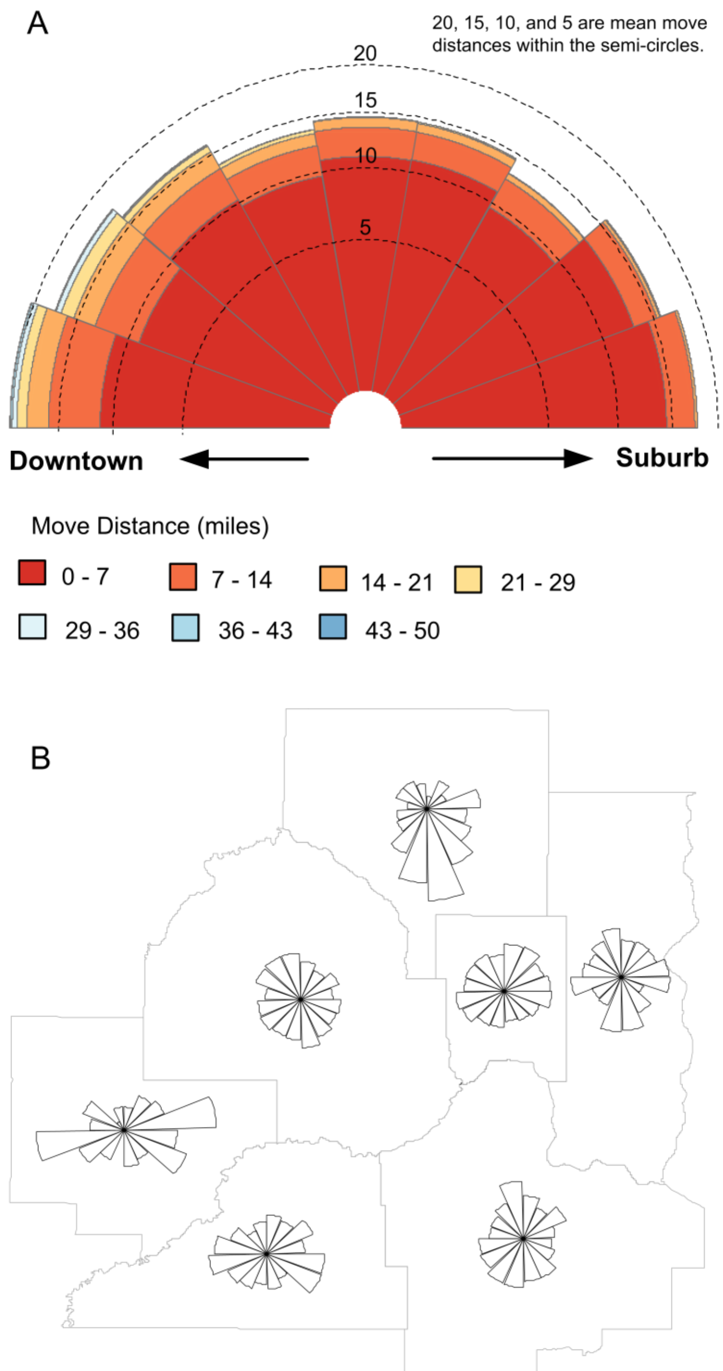

Figure 6.3 Move Direction, Distance, and Volume, 2005-2006. A) Deviation from work-home corridor; and B) Direction and Volume for the Seven Counties in TCMA

There was an interesting transition from moving toward suburbs to moving toward downtown areas, which was not documented in the literature. Before 2005, the mean direction of intraurban migration was around zero degrees, which implies that on average people move outwards from downtown. After 2005, however, the direction was reversed as the mean direction changed from around zero to around 180 degrees or 3.14 in radians, suggesting that slightly more migrants move toward downtown instead of away from it. Of course, this should not be simply interpreted as the rebirth of the urban core, although there was a thread of migration towards the urban center. From the big data analytics of move direction and volume, it was more likely that the suburban areas attracted more rural and exurban residents during the period of strong economic growth in mid-2000s.

Changes of Neighborhood Characteristics after Migration

Examination of the spatiotemporal dynamics of household moves in conjunction with tax data on parcels gives insight into the social mobility of intraurban migration. As discussed earlier, the most frequently cited reason for moving is for better housing. The housing bundle includes not only the housing structure *per se* but also neighborhood features and public services provided by local governments. While elements of better housing are obviously subject to particular households and their different needs, the aggregate pattern of socioeconomic dynamics defined by social mobility and intraurban migration reveals overall trends in housing demand and supply. This research measured socioeconomic dynamics by comparing housing and neighborhood qualities before and after relocation.

In general, intraurban migrants move away from highways, shopping centers, and sporting facilities and move closer to public parks (Table 6.3). The average changes of Euclidean distance to highways, shopping centers, and sports facilities are statistically significant (*t* test). Counterintuitively, people also move away from water (on average they live 256 meters farther from lakes and rivers than their previous dwellings), although this very likely interacts with housing costs (if housing near water is disproportionately more expensive). Only public parks attract people more closely. After migration, people live 24 meters closer to parks on average ($p = 0.02$).

Table 6.3 Change of Neighborhood Characteristics after Migration

Variable	Δ>0	Δ<0	Δ=0	$\bar{\Delta}$	*sd*(Δ)	t-test	[Δ	Δ]
Distance to Water (m)	2481	2201	6	256	1649	0.000	209	303
Distance to Groceries (m)	2599	2087	2	477	3438	0.000	378	575
Distance to Sports Facilities (m)	2724	1962	2	947	5505	0.000	790	1105
Distance to Parks (m)	2248	2436	4	-24	714	0.020	-45	-4
Distance to Highway (m)	2662	2024	2	659	3817	0.000	549	768
White Percentage	2293	1764	631	0.02	0.15	0.000	0.014	0.023
Test Score on Math, English	2029	1686	973	1.16	11.20	0.000	0.836	1.478

Consistent with observations in other metropolitan areas, people moved into neighborhoods with better public elementary schools and higher percentage of whites (Clark, 1976; Komoto, 1994). 58% chose neighborhoods with better elementary schools, 39% worse, and 7% stayed in the same school attendance area. The average gain of white percentage after migration was near two percent in the Twin Cities area, which is not trivial considering at it was about 90% white over the research period.

For the whole TCMA, more people move into newer, larger, and moderately more expensive houses (Table 6.4). During 2005 to 2006, about 66% of intraurban migrants moved to newer houses, 31% to older houses. The average change of the age of migrants' dwellings is 9 years newer. The change of lot size is interestingly consistent with the change of housing age. About 67% intraurban migrants move into houses with larger parcel lots and 32% to smaller. The average increase of lot size is 280 square feet or 26 square meters.

Table 6.4 Change of Housing Structural Characteristics after Migration

Variable	Δ>0	Δ<0	Δ=0	$\bar{\Delta}$	*sd(*Δ)	t-test	[Δ	Δ]
Year Built	2,485	1,177	111	9.02	32.26	0.000	7.99	10.05
Estimated Values	2,376	1,800	11	1,799	148,800	0.434	-2709	6,307
Lot Size (sq feet)	1,290	619	11	280.574	769.8	0.000	246.0	314.9
Property Tax	1,914	1,269	4	235.63	2794.1	0.000	138.6	332.7

Despite the significant gain in the structural age and lot size, the estimated property values before and after migration remain the same. Although the number of households that move into more expensive houses was much more than those moving to cheaper houses, i.e., 57% vs. 43%, the average price change was below $2,000 and insignificantly different from zero. This implies that among those who moved to more expensive houses, most moved to houses just a little bit more expensive, while most migrants who moved to cheaper houses lose a lot in housing value. In other words, the dynamics of financial status was not even. When people moved up on the financial ladder, they just rose a little; when they moved down, they dropped a lot. The property tax paid after migration, however, was significantly more than in previous jurisdictions. The average difference was 236 dollars, which suggests that some residents actually liked to pay more property tax to move into better neighborhoods and to receive better public services.

Conclusion

Intraurban migration is an important aspect of urban and housing studies. This chapter provides an example of using big parcel data to examine intraurban migration. Big data techniques and analytics help derive a relatively exhaustive and fine-scaled delineation of the residential moves within the TCMA. With unmatched spatial accuracy, the extracted migration information helped refine the mathematical models of spatial moves and also provided direct empirical evidence for increasing move distance, directional bias, and the dynamics of urbanization. In addition, exploring the changes of structural and neighborhood characteristics after migration revealed features of social mobility, or how migration and socioeconomic status interact. In summary, big urban data of parcel records offer urban researchers a new opportunity to generate accurate and direct intraurban migration information, to advance our knowledge of housing and migration patterns, and to lay empirical behavioral foundation for modeling the complex housing locational decision process that is critical to understand urban dynamics.

References

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

Adams, J. S., Caruso, D. J., Nordstrand, E. A., & Palm, R. I. (1973). Intraurban Migration. *Annals of the Association of American Geographers, 63*(1), 152-155.

Brown, L. A., & Moore, E. G. (1970). The Intra-Urban Migration Process: A Perspective. *Geografiska Annaler. Series B, Human Geography, 52*(1), 1-13.

Castles, S., Miller, M. J., & Ammendola, G. (2005). The age of migration: International population movements in the modern world. *American Foreign Policy Interests, 27*(6), 537-542.

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

Clark, W. A. V. (1986). *Human migration*. Beverly Hills: Sage Publications.

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

Clark, W. A. V., Deurloo, M. C., & Dieleman, F. M. (2006). Residential mobility and neighbourhood outcomes. *Housing Studies, 21*(3), 323-342.

Clark, W. A. V., Huang, Y., & Withers, S. (2003). Does commuting distance matter? Commuting tolerance and residential change. *Regional Science and Urban Economics, 33*(2), 199-221.

De Jong, G. F., & Roempke Graefe, D. (2008). Family life course transitions and the economic consequences of internal migration. *Population, Space and Place, 14*(4), 267-282.

Fan, C. C. (2002). The Elite, the Natives, and the Outsiders: Migration and Labor Market Segmentation in Urban China. *Annals of the Association of American Geographers, 92*(1), 103-124.

Frey, W. H., & Liaw, K.-L. (2005). Migration Within the United States: Role of Race-Ethnicity. In G. Burtless & J. R. Pack (Eds.), *Brookings-Wharton Papers on Urban Affairs 2005* (pp. 207-248). Washington, D.C.: Brookings Institution Press.

Galster, G. (1988). Residential segregation in American cities: A contrary review. *Population Research and Policy Review, 7*(3), 93-112.

Haynes, K. E., & Fotheringham, A. S. (1984). *Gravity and spatial interaction models*. Beverly Hills: Sage Publications.

Hipp, J. R., & Boessen, A. (2016). The Shape of Mobility: Measuring the Distance Decay Function of Household Mobility. *The Professional Geographer*, 1-13.

Jones, C., Leishman, C., & Watkins, C. (2004). Intra-urban Migration and Housing Submarkets: Theory and evidence. *Housing Studies, 19*(2), 269-283.

Kitchin, R. (2013). Big data and human geography Opportunities, challenges and risks. *Dialogues in human geography, 3*(3), 262-267.

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal, 79*(1), 1-14.

Komoto, C. T. (1994). *Intraurban Migration in the Minneapolis-Saint Paul Metropolitan Area.* Unpublished Dissertation, University of Minnesota, Minneapolis.

Mabogunje, A. L. (1970). Systems Approach to a Theory of Rural-Urban Migration. *Geographical Analysis, 2*(1), 1-18.

Manson, S. M., Sander, H. A., Ghosh, D., Oakes, J. M., Orfield Jr, M. W., Craig, W. J., et al. (2009). Parcel data for research and policy. *Geography Compass, 3*(2), 698-726.

Rae, A., & Singleton, A. (2015). Putting big data in its place: a Regional Studies and Regional Science perspective. *Regional Studies, Regional Science, 2*(1), 1-5.

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

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

Sun, S. (2009). *Intraurban Migration in the Twin Cities Metropolitan Area of Minnesota.* Unpublished Dissertation, University of Minnesota, Minneapolis, MN USA.

Sun, S., & Manson, S. (2012). Intraurban Migration, Neighborhoods, and City Structure. *Urban Geography, 33*(7), 1008-1029.

Sun, S., & Manson, S. (2015). Simple Agents, Complex Emergent City: Agent-Based Modeling of Intraurban Migration. In M. Helbich, J. Jokar Arsanjani & M. Leitner (Eds.), *Computational Approaches for Urban Environments* (Vol. 13, pp. 123-147): Springer International Publishing.

Whitelaw, J. S., & Robinson, S. (1972). A Test for Directional Bias in Intraurban Migration. *New Zealand Geographer, 28*(2), 181-193.

Wu, W., Wang, J., & Dai, T. (2016). The Geography of Cultural Ties and Human Mobility: Big Data in Urban Contexts. *Annals of the American Association of Geographers, 106*(3), 612-630.