Open access • Journal Article • DOI:10.1080/00330124.2016.1157495

The Shape of Mobility: Measuring the Distance Decay Function of Household Mobility

John R. Hipp, Adam Boessen

Institutions: University of California, Irvine, University of Missouri–St. Louis

Published on: 02 Jan 2017 - The Professional Geographer (Routledge)

Topics: Distance decay

Title
The Shape of Mobility: Measuring the Distance Decay Function of Household Mobility

Journal
Professional Geographer, 69(1)

ISSN
0033-0124

Authors
Hipp, JR
Boessen, A

Publication Date
2017-01-02

DOI
10.1080/00330124.2016.1157495

Peer reviewed

The shape of mobility: Measuring the distance decay function of household mobility

John R. Hipp*
Adam Boessen

January 11th, 2016

Post-print. Published in The Professional Geographer 69(1): 32-44

Word count: 7,065
Word count (including references): 8,233
Running Head: "Mobility distance function"

* Department of Criminology, Law and Society and Department of Sociology, University of California, Irvine. Address correspondence to John R. Hipp, Department of Criminology, Law and Society, University of California, Irvine, 3311 Social Ecology II, Irvine, CA 92697; email: john.hipp@UCI.edu. This research is supported in part by NSF grant BCS-0827027

[1] The research in this paper was conducted while the first author was a Special Sworn Status researcher of the U.S. Census Bureau at the Triangle Census Research Data Center. Research results and conclusions expressed are those of the author and do not necessarily reflect the views of the Census Bureau. This paper has been screened to ensure that no confidential data are revealed.

The shape of mobility: Measuring the distance decay function of household mobility

Abstract

A well-known challenge to studies examining the distance of residential mobility patterns is that the estimates are often constrained to only patterns within a particular metro area or between metro areas. Thus, studies are unable to estimate of the entire distance decay functional form. Using a unique dataset on the distance of the most recent move for a large sample of households in 23 metropolitan areas in the U.S. over three waves, we flexibly estimate the distance decay function for the entire sample, as well as for a series of subpopulations based on key demographic information.

Keywords: residential mobility, distance decay, gravity model,

Bio

John R. Hipp is a Professor in the departments of Criminology, Law and Society, and Sociology, at the University of California Irvine. His research interests focus on how neighborhoods change over time, how that change both affects and is affected by neighborhood crime, and the role networks and institutions play in that change. He approaches these questions using quantitative methods as well as social network analysis. He has published substantive work in such journals as *American Sociological Review, Criminology, Social Forces*, *Social Problems, Mobilization, City & Community, Urban Studies* and *Journal of Urban Affairs*. He has published methodological work in such journals as *Sociological Methodology*, *Psychological Methods*, and *Structural Equation Modeling*.

Adam Boessen is an Assistant Professor in the department of Criminology and Criminal Justice at the University of Missouri St. Louis. His primary research interests include neighborhoods and crime, geography and space, and social networks.

The shape of mobility: Measuring the distance decay function of household mobility

Geographers and demographers have long been interested in the question of how far households travel when engaging in residential mobility (e.g. see Ravenstein 1885). Studies have explored this question using different data structures, and different data analytic techniques, providing key insights (e.g., Anderson 1956; Clark 1986; Galle and Taeuber 1966). One key insight that the literature has established quite clearly is that households are more likely to move shorter distances rather than longer distances (La Gory and Pipkin 1981). This distance decay pattern is arguably explained as a result of cost (i.e., short distances cost less) and familiarity with the area. Studies have found this effect when focusing on the likelihood of moves within vs. between metropolitan areas (Long 1988), within a particular metropolitan region (Brown, Horton, and Wittick 1970; Clark 1976), and between metropolitan areas (Anderson 1956; Galle and Taeuber 1966). Nonetheless, studies are frequently limited to information on only intra-metropolitan moves, or inter-metropolitan moves, which precludes estimating the entire distance decay function to capture the distance of moves over all possible moves.

A well-known challenge to studies examining the distance of residential mobility patterns is that their estimate of the distance decay function is truncated. This can occur in three ways: 1) when capturing moves within the metropolitan area, these studies are unable to estimate the tail of the general residential mobility function given that they, by definition, have no information on such longer moves, 2) studies that focus on moves between metropolitan areas do not have information on moves within their metro regions, and thus are missing information on shorter moves, and 3) there is a natural truncation in the estimated functional form given the specific spatial distribution of where housing units are located in the city or metropolitan area (i.e., some housing footprints are constrained by large bodies of water or mountainous areas). If the distance decay function is nonlinear, the estimated functional form for short distance moves (within a metropolitan area) may appear quite different from the estimated functional form for long distance moves (between metropolitan areas). One perspective might suggest that there really are two distinct functions, whereas another perspective might suggest that it could be because the function is not being estimated on the entire distribution of possible moves (both intra- and inter-metropolitan areas). This distinction is quite notable, given that in the 2010 United States Census, 52% of moves where intra-country, 42% inter-county, and 6% from abroad (Ihrke and Faber 2012). A substantial proportion of the distance decay distribution is missing by only focusing on inter or intra patterns, and thus there is a need to explore this distance decay pattern without truncating the distribution to only inter or intra mobility patterns.

Beyond the question of how far households are likely to move, is the question of whether certain types of households are more likely to move farther distances than others. That is, are there differences by socio-economic status, or by social demographic characteristics, in the average distance that residents move? These differences can manifest across average distance moved, but these differences can also manifest across the entire range of moves by distance for different groups. For example, if certain groups are less likely to make short distance moves, but do not differ in their likelihood to make long distance moves, this pattern will be reflected when viewing the entire range of moves by seeing a difference at shorter moves that evaporates at longer distance moves.

As Niedomysl and Fransson (2014) note: "…in the vast majority of countries, only migration flow data between relatively large administrative areas are available, making it virtually impossible to use actual distance as a migration-defining criterion." The present study is therefore able to make an important contribution by using a sample of households in 23 different metropolitan areas over three time points in the United States to estimate the distance decay function of the most recent move of the household based on small geographic units capturing the origination and destination location (zip codes and tracts). We move beyond prior research with a panoramic view of migration distances that includes inter and intra mobility patterns. This allows us to: 1) estimate the distance decay functional form of all moves; 2) estimate the proportion of moves that are within the same neighborhood; 3) assess the differences in average move distance across various socio-demographic characteristics; and 4) assess the differences in the range of distances that various socio-demographic groups move. Whereas some recent work has viewed all residential mobility in Sweden with precise geographic data (Niedomysl and Fransson 2014), our work builds on this research by estimating models for the United States. With data from the U.S., we can contribute new insight to mobility distances for even longer potential moves, for a substantially larger country than prior work (i.e., Sweden), and a country with considerably more diversity (i.e., race/ethnicity). Moreover, prior research often only focuses on the average distance moved, and another contribution of our project is assessing the entire distance distribution of moves for different subgroups. We are aware of no research to date that has examined any of these patterns, and these estimates are informative given the long history of data challenges in this area. The results may also be of interest to scholars wishing to build agent based simulations and those interested the formation of segregation patterns.

Distance and Residential Mobility

Given that distance is often found to be one of the most important factors for mobility patterns, scholars have studied the physical distance of moves by households in several fashions, and we now describe this literature.

Mobility distance function

Inter-metropolitan mobility

Numerous studies have focused on the aggregated flows of households in an effort to estimate the distance of residential moves. This literature has largely built on gravity flow models: the notion that distance has important effects inhibiting mobility (e.g. see Zipf 1949). These studies have focused on both inter-metropolitan area mobility flows between large administrative units, as well as intra-urban mobility flows (Haynes and Fotheringham 1984). These studies have consistently shown the importance of the gravity flow model: that is, households are more likely to move to closer destinations rather than further ones. For example, one study looked at aggregate flows between metropolitan areas in the 1930's and detected such an effect (Anderson 1956). Studies have also viewed mobility flows between states (Jun and Chang 1986; Plane and Mulligan 1997) and inter-metro area mobility flows (Galle and Taeuber 1966). Another study focusing on moves between metropolitan areas attempted a hierarchical classification of places as a result of these movement patterns (Brown, Odland, and Golledge 1970).

Although these studies provide important information on patterns of long distance moves, there are certain limitations. One limitation is that these studies only capture long distance moves, and not short distance moves. It is well-known that the motivation typically differs for short- versus long-distance moves. Scholars have long suggested that long distance moves—migration—are presumed to be in response to employment opportunities; furthermore, a recent study of county to county flows suggested that long-distance moves can also conform to stage of life course: whereas young adults tend to move to larger metropolitan areas with job opportunities, older persons nearing retirement are more likely to move to smaller micropolitan areas (Plane and Jurjevich 2009). Research has even suggested that the spatial dispersion of persons' network ties can impact the likelihood of moving further distances from where one was born (Viry 2012). For all these reasons, the assumption that the functional form is similar between long- and short-distance moves may not be reasonable. In fact, a recent study in Italy found differences in long-distance versus short-distance moves, concluding that whereas short distance moves reflect an equilibrium model, long distance moves reflect a disequilibrium model (Biagi, Faggian, and McCann 2011). A second limitation is that such research based on aggregate flows cannot account for the individual motivations and determinants that lead to such moves (Greenwood, Mueser, Plane, and Schlottmann 1991). To the extent that such broader patterns are confounded by systematic individual characteristics, the estimated distance function will be impacted.

A third issue with studying migration flows capturing long distance moves between geographic units is that these studies are often constrained to large areal units. A consequence is that there is considerable uncertainty regarding where exactly the migrant began and ended the move, and therefore estimating distance is quite challenging (Bell, Blake, Boyle, Duke-Williams, Rees, Stillwell, and Hugo 2002). This uncertainty is based on the size of the units (Boyle and Flowerdew 1997). Thus, simply using the population weighted centroid of the units for computing distance obscures the actual level of uncertainty in these measurements. A recent study using data in the U.K. showed that aggregating to successively larger geographic units for the initial and destination location results in more intra-zonal moves and hence greater error in the measure of distance (Stillwell and Thomas 2015). The large size of these areal units will be even more problematic when measuring the shorter distance moves as the signal to noise ratio will become unacceptably low given that the spatial uncertainty of the unit is so large relative to the distance of the move (Boyle and Flowerdew 1997). A fourth issue is that the location of the macro geographic unit within the larger context of the country can have important effects, both on the expected average distance as well as the direction of moves (Eldridge and Jones 1991).

Intra-metropolitan mobility

Another body of literature has focused on the mobility decisions of individual households. This literature has generally focused on mobility flows within a particular city or metropolitan area. An advantage of studies that focus on intra-metropolitan moves based on household-level data is that they can take into account household-level characteristics that come into play in the mobility decision (Quigley and Weinberg 1977). Certain characteristics can impact the distance residents move, which can have consequences for the observed move distances. For example, one study focused on moves within Milwaukee tested the effect of economic opportunities on geographic moves and found that households tended to move to neighborhoods with a similar socio-economic status of their originating neighborhood; nonetheless, there was still notable evidence of a distance decay function (Clark 1976). These characteristics are considered important enough that in instances in which such information is not available, one proposed solution is to use information from different geographical areas, or different time periods, to refine the mobility estimates that are obtained (Rogers, Raymer, and Willekens 2003).

Although a strength of intra-metropolitan studies is that they can estimate more precisely the distance decay function of moves, particularly shorter distance moves, they nonetheless are estimating a constrained function. In part, this distance decay function is constrained only to intra-metropolitan moves, and cannot estimate the distance decay function of longer moves. Instead, such studies are forced to extrapolate beyond the data to estimate the form of the distance decay function. On the other end of the distribution, this pattern may have also consequences for shorter moves as one study examined the distance decay of intra-metropolitan area moves in Cedar Rapids, IA of about 200 households and found that the model tended to over-allocate persons to nearby areal units (Brown, Horton, and Wittick 1970).

These studies are also constrained to the actual size and shape of the metro area and the specific peculiarities of how the population is distributed spatially (Clark 1986). For instance, the presence of bodies of water, or mountains, that create open spaces without any population will impact the particular parametric distance decay function that is estimated. An important consequence is that the specific shape of metropolitan areas (e.g., the distribution of owners and renters) can impact distance decay functions, sometimes in manners that can be predicted analytically (Taylor 1971). Scholars have argued that this tendency to move to closer neighborhoods is in part due to spatial awareness: persons tend to move to areas with which they are familiar, often due to daily travel patterns (Adams 1969). That is, the spatial footprint of the population in an area constrains the possible mobility choices of households, which then will strongly affect the possible shape of the distance decay function that is estimated for households.

Distance decay functions

Although it is challenging to estimate the actual distance decay function that captures mobility, studies have provided at least some empirical evidence showing the relative composition of local versus long distance moves. For example, a study used county centroids to compute mobility distance in the U.S. and found that over 70% of moves were less than 50 kilometers (31 miles) (Long, Tucker, and Urton 1988b). Other research used self-report data regarding the distance of the most recent move in the U.S. and found that 23% moved less than 5 kilometers, although nearly 30% moved at least 80 kilometers (Long, Tucker, and Urton 1988b). This same study also found that more frequent movers self-reported farther distance moves on average than those who move less often (Long, Tucker, and Urton 1988a). An international comparative study of small geographic units found the median distance moved in Australia to be 5.1km compared to a median distance moved of 6.3km in Britain (Bell, Blake, Boyle, Duke-Williams, Rees, Stillwell, and Hugo 2002).

There are in fact many different possible mathematical forms that can characterize the distance decay of mobility distances (Taylor 1975). The actual functional forms that are observed can be due to a mix of social factors, including the spatial patterning of available housing, the spatial pattern of neighborhoods based on socio-economic status and the level of socio-economic segregation, as well as other household-level decision processes. Morrill and Pitts (1967) discussed four for various types of flows: Pareto, exponential, lognormal, and Pareto-exponential. They argued that existing evidence suggested that exponential functions seemed more appropriate for migrations and marriage distances. They then displayed examples of fitted curves for a handful of cities (Morrill and Pitts 1967). In the present study we flexibly estimate the distance decay functional form using multiple polynomials, rather than attempting to adjudicate between these particular parametric forms. This flexible approach allows us to more precisely capture the functional form, rather than imposing an overly strict mathematical function on the data.

Although the existing evidence of inter-regional mobility flows and intra-metropolitan studies finds that distance exhibits a strong diminishing effect on mobility probabilities (Clark 1986), this literature has typically not been able to estimate this entire distance decay function. Studies focusing on intra-metropolitan flows are only able to estimate this functional form at the shorter end of the mobility distance scale. These studies are forced to extrapolate beyond their data to assess the effect of longer distances on mobility flows. And studies looking at inter-metropolitan area flows are able to get estimates of the long tail of such mobility flows, but again must extrapolate to shorter moves. It is an open question whether the functional form estimated for short moves would in fact be appropriate for long moves. Likewise, it is not clear that the functional form estimated for long moves would in fact fit for short moves, and in fact some recent evidence suggests that such estimates will not be particularly accurate (Niedomysl and Fransson 2014). Thus, it is likely inappropriate to extrapolate from the functional form for short distance moves to that for long distance moves, and vice versa.

Differences in movement distances across socio-demographic groups

Beyond the question of the distances households typically move is the question of whether there are differences in mobility distances among subgroups. Studies of intra-metropolitan moves have detected differences in distance moved based on the characteristics of the household.[1] Whether the same differences would be detected for longer-distance moves is an open question. We now discuss several of these characteristics.

One important characteristic that affects the average distance moved is the socio-economic status of the household. Those with more economic resources are typically better able to undertake longer distance moves. In part, this is because of greater job opportunities due to better educational training. A study of Providence, RI, found that those from higher socio-

[1] Studies have also extended the gravity flow model by including various sociocultural characteristics of the destination as predictors of direction of mobility for inter-state flows (Herting, Grusky, and Van Rompaey 1997). Whereas the early studies focused almost exclusively on the gravity flow model and the importance of physical distance, later work increasingly incorporated other important characteristics of metropolitan areas that might influence migration destinations (Ferguson and Kanaroglou 1995).

economic neighborhoods were more likely to move longer distances (Simmons 1968). A study in Germany found that more highly educated persons were more likely to make long distance moves than those with lower education (Hunt 2004). And research in the U.K. found that those with more income and higher education tended to move longer distances (Thomas, Stillwell, and Gould 2015). A study of Sweden using relatively accurate geographic information also found that those with higher levels of education tended to move farther; however, a somewhat surprising result was that this same study found that those with higher income actually made *fewer* long distance moves (Niedomysl and Fransson 2014). Research has suggested that even commuting patterns tend to be impacted by the spatial patterns of neighborhoods based on socio-economic status (Kipnis and Mansfeld 1986; Simmons 1968).

Another socio-demographic characteristic that might impact the distance a household moves is race and ethnicity. Even though there tends to be a strong correlation between the race of the household and socio-economic status, there is also reason to suspect that differences in move distance across racial groups may be due to constrained choices—beyond economic constraints—as racial minorities often have limited ability to move into many neighborhoods. This occurs due to steering by real estate agents and due to discrimination faced by racial/ethnic minorities in the housing market (Crowder and South 2005; Fischer and Massey 2004; Massey and Denton 1993). Indeed, there is some evidence that blacks typically make shorter moves (Clark 1986). Similarly, other research finds that whites are more likely to make long-distance moves than racial minorities (La Gory and Pipkin 1981; McCarthy, Valenstein, and Blow 2007). There is less evidence regarding the distances Latinos typically move compared to other groups.

The distance of moves may also differ between owners and renters. In part this may occur because renters often represent younger residents with lower socio-economic status.

However, this may also occur because renters are less attached to the neighborhood than owners (Lee, Campbell, and Miller 1991; Mesch and Manor 1998) and less satisfied (Hipp 2009; Hipp 2010; Lu 1999), both of which will increase mobility. This suggests that renters will be more likely to make short distance moves given their lack of attachment to, or satisfaction with, the neighborhood, even if they do not differ in likelihood for making long-distance moves. Indeed, evidence suggests that renters are more likely to move than owners, and that they move shorter distances (Clark 1986).

Other demographic characteristics such as age and family structure may be important. There is strong evidence that the relative frequency of moves is related to age, and this even differs based on the region of the country (Pandit 2000). Given the robust evidence of a nonlinear relationship with age in which the oldest and the youngest are the most likely to move in general (Crowder 2001; South and Crowder 1997), it may be that this tendency also results in more short distance moves for these age groups. As some support for this, a study of divorcees in 2000 in the Netherlands found that younger persons moved shorter distances on average compared to older persons (Feijten and Van Ham 2007: 640). However, evidence from Sweden (Niedomysl and Fransson 2014) and the U.K. (Thomas, Stillwell, and Gould 2015) finds that those who are older are *less* likely to make long distance moves. Similarly, the evidence that households with children are less likely to move in general (Deane 1990; South and Crowder 1998) may also translate into the relative distance of such moves. A study of households in Sweden did find that households with children tended to make fewer long distance moves (Niedomysl and Fransson 2014). And a study of patients treated in the Veterans Affairs (VA) health system found that those who are married moved farther (McCarthy, Valenstein, and Blow 2007).

We are able to explore these questions here given that we have a sample of households from a large number of metropolitan areas, and we have information on the distance of their most recent move no matter where they were moving from in the entire United States. We therefore explore: 1) the average and median distance of moves by households of various demographic subcategories; 2) the distribution of residential moves based on distance for the entire sample, as well as households of various demographic subcategories; 3) multivariate analyses assessing the relative contribution to distance of recent moves for various demographic subcategories. We describe the data next.

Data and Methods

Data

The American Housing Survey (AHS) conducts surveys of approximately 4,000 housing units from each of a large number of metropolitan areas across the U.S in various years. Every two years the AHS surveys a subset of the metropolitan areas: as a result, a particular metropolitan area is surveyed approximately every four years. Because of this variability in the actual year of the survey, we are sometimes combining metropolitan areas from slightly different years. That is, whereas the "waves" are labeled 1987, 1991, and 1995, these "waves" actually contain the data for the nearest year in which a particular metropolitan area was surveyed. For instance, whereas in the wave 1987 some of the metropolitan areas were actually surveyed that year, some of the metropolitan areas were actually surveyed in 1985. We have a very large sample of households in 23 metropolitan areas in 1987, 22 in 1991, and 23 in 1995. Pooling across these 68 metropolitan years, we have 66,383 household moves.[2]

By using special access to a Census Research Data Center, we were able to place AHS residents into 1980 census tract boundaries. Thus, we knew the current tract of residence. The AHS asks respondents to report on the zip code of their previous residence. Thus, we have geographic "containers" of the origin and destination of the most recent move for each respondent in the sample. These are not ideal geographic containers, and therefore pose specific challenges. First, zip codes are not ideal geographic containers given that they were created by the Postal Service for the express goal of delivering mail, and therefore do not necessarily map onto the concept of "neighborhood". Furthermore, zip codes can change boundaries quite readily, with minimal documentation of the changes over time. As a best approximation of the zip code boundaries over the time period of our study, we used zip code boundaries from 1991.[3] From these boundary files, we computed the latitude/longitude of the center point of each of these zip codes. From the same source, we obtained 1980 tract boundaries, and computed the latitude/longitude of the center point of each of these tracts. In 2000, the median census tract in the U.S. was just under 2 square miles, whereas the median zip code was 38 square miles.

We also computed estimates of the number of moves that occurred within the "same neighborhood" (using the tract as the "neighborhood"). A challenge is that zip codes and tracts only partially overlap. We geographically overlaid the 1991 zip code boundaries with the 1980 tract boundaries, and if there is no overlap of the zip code and tract boundaries, then we know the household moved *into* the tract (and not within the same tract). If there is partial overlap in the

[2] The metropolitan areas are: Anaheim, Baltimore, Buffalo, Chicago, Cleveland, Dallas, Denver, Fort Worth, Houston, Indianapolis, Los Angeles, Milwaukee, Minneapolis, Oakland, Philadelphia, Phoenix, Portland OR, Riverside-San Bernardino, Sacramento, Saint Louis, San Diego, San Francisco, Seattle, Tampa, Washington DC.
[3] These were obtained from the MABLE/Geocore website located at the Missouri Census Data Center website (http://mcdc2.missouri.edu/websas/geocorr90.shtml).

boundaries we cannot be certain if mobility occurred within the same tract. We adopted two approaches to deal with this uncertainty. First, we created an upper bound estimate by coding the resident as moving within the same neighborhood if there was *any* overlap at all between the zip code boundary and the tract boundary (no matter how small). Second, we assumed that a person from a zip code that partially overlaps with the current tract of residence has a uniform random chance of where they lived in the zip code. Therefore, if 10% of the previous zip code overlapped with the current tract, we assumed that the household had a 10% chance of moving from the portion of the zip code contained within the tract (and a 90% chance of having moved from a different tract). If residential moves in fact exhibit a distance decay, this uniform distribution assumption is not quite accurate and would lead to a small underestimate of the probability of moving within the same tract. Nonetheless, these two approaches give us bounds on the percentage of households moving within the same neighborhood.

Dependent Variables

We computed the distance in miles between the centroid of the tract of residence and the zip code that the household reported moving from. We log transformed this value.

Independent variables

In the initial analyses in which the outcome was logged distance of the previous move, we created a series of standard demographic measures capturing information on the household head or the household in general to assess their relationship with move distance. We created indicators of whether the household head is African American, Latino, Asian, or other race (with white as the reference category). We computed a measure of years of education of the household head, and a measure of household income. We capture the distinction between owners and renters with an indicator of owners. We created measures of age of household head, and age squared, to capture possible nonlinearities in distance moved over the life course. We capture household composition with indicators of currently married, widowed, or divorced (with never married as the reference category), and indicators of whether the household has children less than school aged (less than 6 years of age, but no school-aged children), or whether the household has school aged children (aged 6 to 18 years), with the reference households without children under the age of 18.

Methods

In all models, the outcome measure is logged distance of the most recent move. The models are estimated as ordinary least squares regressions. We begin by estimating multivariate models that capture the difference in move distance based on the demographic characteristics of the household, which are the demographic variables previously described.

In the second set of models, we estimate the distance decay function of the most recent move. In these models that capture the distance distribution, we first ordered the sample from shortest to longest distance, and then created an integer indicator ranging from 1 to the sample size based on this ordering (thus, the household with the shortest distance move is coded 1, the household with the second shortest distance move is coded 2, on up to the household with the longest distance move which is coded to the sample size value). We also created various polynomials of these integer indicators to capture nonlinearities in the distance function. The predictor variables are the integer indicator showing the relative distance of the move for a particular household, and as many polynomials of this as are necessary to reasonably capture this distance function (polynomials raised to the 5^{th} or 6^{th} power captured the patterns). These models do an extremely good job of explaining this distribution. For example, the R^2 for the model of the full sample was .995. Thus, we are effectively capturing this functional form of this distance distribution. By then plotting these polynomials we can characterize the functional form of the distance decay over the entire sample. The models for the various subsamples also do an excellent job capturing their distance decays with R^2's ranging between .992 and .997. We used the same procedure for each of the sub-samples on which we computed the distance distribution as we did for the entire sample.

Results

We first present the summary statistics for the distance of moves for the sample and the sub-populations of interest in the sample. As seen in Table 1, the median distance of moves in this sample is 4.4 miles, whereas the mean distance is 114 miles; this of course implies the unsurprising considerable skew given that some moves can be for extremely long distances, whereas the bulk of moves are for very short distances (see Figure 1 for the percentage of moves at various distances). For those in poverty, moves are about 25% shorter distance (median 3.3 miles). Whites tend to move about 50% longer distances (median is 4.9 miles) compared to African Americans and Latinos (a median of about 3.3 miles). Owners move much longer distances (about 75% longer) compared to renters: 6.4 miles compared to 3.7 miles for the median values. The average distance of moves for those with a bachelor's degree are 85% longer than for those with only a high school degree.

<<<Table 1 about here>>>

<<<Figure 1 about here>>>

Mobility distance function

In this same table we also see the estimates of the percentage of moves within the same neighborhood for these various subpopulations. As described earlier, given the geographic uncertainty of our data, we provide upper and lower bound estimates. For the complete sample, in their last move between 16% and 21.5% moved within the same tract. Since it is likely that the true values are closer to the lower bound, we compare groups based on these values. Thus, the percentage moving within the same tract is higher for households that are in poverty (18% lower bound estimate), black (18.6%), Latino (17.3%), or with less than a high school degree (20%). In contrast, whites, owners, and those with higher levels of education are less likely to move within the same tract.

Multivariate results

We next assessed the partial correlation effects of these various socio-demographic characteristics on the distance of moves, holding constant these demographic characteristics. The results of this model are shown in Table 2. Given that the outcome is logged distance, we can interpret these coefficients in terms of percentage changes in the distance of moves. Holding these demographic characteristics constant, an African American household moves about 50% shorter distance than does a white household (the reference category). Whites in general move the longest distances, as Latinos move 32% shorter distance than whites, Asians move 25% shorter distance, and other race households move 28% shorter distance. Thus, these racial differences are observed even when controlling for socioeconomic differences across groups.

<<<Table 2 about here>>>

Higher socio-economic status (SES) households move greater distances, holding constant these characteristics. Each additional year of education increases the distance moved about 8%, whereas each additional $1000 in income increases the distance moved almost 10%. Whereas prior studies nearly always find that renters move more frequently than owners, we also see evidence here that renters move *shorter distances* than owners (on average, about 7% shorter distances, holding constant these other characteristics).

The relationship between age and the distance of moves is u-shaped. At the youngest and oldest ages, households are most likely to move the longest distances. The inflection point is at 37 years of age, suggesting that households with a 37 year old household head move the shortest distances, on average. For example, a 20 year old, on average, moves about 6% farther than a 37 year old, holding constant these other household characteristics. A 60 year old, on average, moves about 12% farther than a 37 year old, holding constant these other household characteristics.

We also see important differences for household structure. Whereas prior studies have frequently shown that single person households tend to be quite mobile, we see that when they do move they tend to move shorter distances. For example, married households tend to move much farther than single person households (about 30% farther, on average, holding constant these other demographic characteristics), and widowed households also move farther than single person households (about 8% farther, on average). Households with children move shorter differences than those without children. Households with school-aged children (aged 6 to 18 years) move about 7% shorter distances on average than households without children, and households with very young children (less than 6 years of age) move about 15% shorter distances than households without children.

Exploring the functional form of the distance of moves

Whereas our results to this point have compared the central tendency of moves among various groups, we next ask about the functional form of the whole range of distances moved, as well as possible differences across groups. We plot each of these distance decay functions based on logged distance given that it more closely characterizes the pattern compared to raw distance. For the complete sample, we see in Figure 2 that a logged relationship would capture the middle range of the distance of moves, but there are bends at the two ends of the distribution. This figure shows shorter to longer distance moves, and the cumulative proportion of the sample that has moved a particular distance is plotted on the x axis whereas logged distance is plotted on the y axis. That is, given the slope in the middle of this figure, there are fewer very short distance movers and fewer very long distance movers compared to the proportion in the middle portion of this distribution. Thus, based on these estimates, the 10th percentile of move distances is 0.95 miles, the 20th percentile is 1.68 miles, the 30th percentile is 2.33 miles, the 40th percentile is 3.2 miles, and the median is 4.45 miles. At the other end of the distribution, the 80th percentile is 11.75 miles and the 90th percentile is 37.7 miles. Thus, only about 8.5% of the sample moved 50 or more miles, 5.7% moved 100 or more miles, and 3.3% moved 200 or more miles.

<<<Figure 2 about here>>>

We next focus on the distance distribution of moves for various subpopulations of our sample. We illustrate the distance distribution for those in poverty by overlaying it on the distribution for the entire sample in Figure 2. As expected, we see that those in poverty move shorter distances than the complete sample (which includes households both in poverty and not in poverty). At virtually all points of the distribution, households in poverty move shorter distances than those of the entire sample. For example, at the 30th percentile, the typical household in the sample moved 37% farther than a household in poverty (2.33 to 1.7 miles), and this relative gap remains constant up through the 70th percentile. However, this gap widens for the longest moves: at the 80th percentile this difference is 56% farther, and at the 90th percentile it is 127% farther. It appears households in poverty are particularly unlikely to make the longest moves---those that are inter-regional.

Turning to the differences by race/ethnicity, we see that Latinos and African Americans typically move shorter distances than whites, although this pattern changes at the longest distances. Figure 3 compares the distance distributions of Latinos, African Americans and Whites and shows that at shorter distances African Americans and Latinos consistently move much shorter distances than whites. At the 30th percentile, a white household moves 700% farther than does an African American household and 650% farther than a Latino household. At the 50th percentile this gap is about 600% for whites compared to both groups, and it is approximately 275% at the 70th percentile. By the 90th percentile this gap has narrowed such that whites move 142% farther than African Americans and 32% farther than Latinos. But at the longest distances these patterns reverse, and blacks move farther than whites for the 3% longest distance moves and Latinos move farther than whites for the 8% longest distance moves. This is a pattern that can only be detected by analyzing the entire distance spectrum of moves by these groups.

<<<Figure 3 about here>>>

When comparing households based on the presence of children, there are few differences for short distance moves (see Figure 4). It appears that an equal proportion of households with very young children, school aged children, or no children, move similar distances among the 67% shortest distance moves. However, there are differences for longer moves: those with very young children move considerably shorter distances than those with school aged children or no children. By the 80th percentile, households with children 6 to 18 years old move 58% farther than households with children 5 and younger, and this gap grows to 134% at the 90th percentile.

Households without children move 58% farther than households with children 5 and younger at the 80th percentile, and over twice as far at the 90th percentile.

<<<Figure 4 about here>>>

There are also consistent differences in the distance of moves for owners and renters at all distances in the distance distribution, as seen in Figure 5. Although owners are less likely than renters to move, when they do so, they move longer distances. Owners tend to move 60 to 80% farther than renters at all points in the distribution, with the only narrowing of the gap occurring at the 5% longest moves.

<<<Figure 5 about here>>>

Conclusion

We have extended the literature on the distance traveled by households when changing residence. Although questions regarding migration distance stem from early work in geography, progress in this area has arguably been slow due to little available data (Niedomysl and Fransson 2014). Whereas prior literature often is constrained to assessing the distance decay function either on short moves only (within a metro area) or on long moves only (inter-regional moves), we were able to estimate with a unique dataset the distance of household moves over the complete population of moves within the United States for households in 23 metropolitan areas. We were also able to compare the distances of moves among various demographic subpopulations.

One important contribution of this study is that we were able to estimate the distance decay function over all moves—both intra- and inter-metropolitan moves. This allowed us to capture the functional form of this decay function over the entire range of possible moves; although this has been studied in Sweden (Niedomysl and Fransson 2014), it is not been studied across a country of the large geographic size of the U.S. Over most distances, a log-linear representation appeared to capture this distance decay function. Nonetheless, it is interesting to note that the distance decay function showed nonlinearity at the longest distances, as there were fewer observations among long distance movers than would be expected based on the rest of the distribution. This is likely in part due to the fact that we did not have geographic information on immigrants. A benefit of our approach that provides this accurate distance decay estimate is that it can be used by studies that wish to build a model of mobility patterns based on agent-based simulations (Sun and Manson 2010). A more accurate estimate of this functional form will improve the predictions from such simulation models.

Although some of our results when assessing the distance decay function of mobility for various subpopulations confirmed those in the existing literature, there were nonetheless some novel findings. For example, regarding the socioeconomic status of the household, we found that those with higher levels of education moved farther distances, which is consistent with the extant literature. However, we also found that those with higher household income also moved longer distances, which is opposite of recent evidence from Sweden that also used small geographic containers of origin and destination (Niedomysl and Fransson 2014). And when we studied the entire range of moves, we found that households in poverty move shorter distances than other households at virtually all distances of moves: short, medium, and long distances. As a result, households in poverty, household heads with less than a high school degree, and Latino and or African American household heads, were the most likely to make moves within the same neighborhood. The fact that those in poverty are very likely to move within the same neighborhood is consistent with the idea of such residents being trapped in a cycle of poverty (Wilson 1987). Future research might explore these differences between the U.S. and Sweden for poverty, and one possibility is that cities in the U.S. might have concentrated poverty in central cities, whereas European cities may concentrate poverty in the outskirts of city implying differences in distance of residential moves between the two counties (e.g., see Brueckner, Thisse, and Yves Zenou 1999).

Our relatively novel contribution of estimating the entire distribution of moves across subgroups allowed us to detect differences among groups between short- and long-distance moves. First, whereas households with children move shorter distances, on average, than other households, we found that the presence of children made *no difference* in the distance of moves for short distance moves. It was only for the longest distance moves that we detected that households with very young children (less than school age) are less likely to make these long distance moves. Second, whereas we found that renters move shorter distances than owners, on average, which is consistent with extant literature (Clark 1986), we also detected that this gap between renters and owners is only present for shorter moves and evaporates for the very longest moves. The fact that owners are less likely to make shorter moves is likely consistent with the evidence that they are typically more satisfied with their neighborhood (Hipp 2009; Lu 1999) as well as more attached to the neighborhood (Lee, Campbell, and Miller 1991; Mesch and Manor 1998) , which likely limits the number of short distance moves they make. Third, the differences in mobility distances among households based on race/ethnicity were particularly striking: whereas whites move longer distances on average than African Americans and Latinos, which is consistent with the findings of existing literature, we found that this pattern actually reverses for extremely long distance moves. Such an effect is obscured when simply comparing average distances of moves. Nonetheless, this was a strong effect that we detected here, and to our knowledge has not been noted in the existing literature.

We acknowledge some limitations of this study. Because we had no information on the prior location of residence for those who emigrated from another country, we were unable to estimate the extreme ends of the distance decay distribution. As of the 2010 Census, these movers from abroad represented approximately 6% of the distribution of movers, so we suspect their impact may be relatively small. We also did not account for the recency of the move; instead, we only utilized information on the most recently reported move. Furthermore, as discussed, the modest imprecision of the geographic containers that the household lives in and came from (tracts and zip codes) introduce some measurement error into our estimates of the shortest distance end of this functional form. Nonetheless, recent research comparing centroids of small geographic containers for municipalities to actual distances with a large sample suggest little to no differences in regards to the accuracy of the estimates (Niedomysl, Ernstson, and Fransson 2015).

In conclusion, understanding the complete distance decay function for households, as well as subsets of households, provides key information for understanding various demographic processes. This distance decay function can vary across contexts (Eldridge and Jones 1991), and understanding the specific shape of this functional form for specific subpopulations provides unique information regarding the character of mobility for these groups beyond measures assuming proportional differences in mobility distance. We were able to detect average distances across groups when viewing the entire range of moves—both intra- and inter-metropolitan moves. We were also able to detect differences among subgroups that sometimes only occurred among the shortest moves, and other differences that only occurred among the longest moves. Such information helps in understanding better the general patterns of mobility distances among subgroups.

References

Adams, John S. 1969. "Directional bias in intra-urban migration." *Economic Geography*:302-323.

Anderson, Theodore R. 1956. "Intermetropolitan Migration: A Correlation Analysis." *American Journal of Sociology* 61:459-462.

Bell, M., M. Blake, P. Boyle, O. Duke-Williams, P. Rees, J. Stillwell, and G. Hugo. 2002. "Cross-national comparison of internal migration: issues and measures." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165:435-464.

Bell, Martin, Marcus Blake, Paul Boyle, Oliver Duke-Williams, Philip Rees, John Stillwell, and Graeme Hugo. 2002. "Cross-national comparison of internal migration: issues and measures." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165:435-464.

Biagi, Bianca, Alessandra Faggian, and Philip McCann. 2011. "Long and Short Distance Migration in Italy: The Role of Economic, Social and Environmental Characteristics." *Spatial Economic Analysis* 6:111-131.

Boyle, Paul J and Robin Flowerdew. 1997. "Improving distance estimates between areal units in migration models." *Geographical Analysis* 29:93-107.

Brown, Lawrence A, John Odland, and Reginald G Golledge. 1970. "Migration, functional distance, and the urban hierarchy." *Economic Geography*:472-485.

Brown, Lawrence A., Frank E. Horton, and Robert I. Wittick. 1970. "On Place Utility and the Normative Allocation of Intra-Urban Migrants." *Demography* 7:175-183.

Brueckner, Jan K, Jacques-Francois Thisse, and Yves Zenou. 1999. "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory." *European Economic Review* 43 (1):91-107.

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

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

Crowder, Kyle D. 2001. "Racial Stratification in the Actuation of Mobility Expectations: Microlevel Impacts of Racially Restrictive Housing Markets." *Social Forces* 79:1377-1396.

Crowder, Kyle D. and Scott J. South. 2005. "Race, Class, and Changing Patterns of Migration between Poor and Nonpoor Neighborhoods." *American Journal of Sociology* 110:1715-1763.

Deane, Glenn D. 1990. "Mobility and Adjustments: Paths to the Resolution of Residential Stress." *Demography* 27:65-79.

Eldridge, J Douglas and John Paul Jones. 1991. "Warped space: A geography of distance decay*." *The Professional Geographer* 43:500-511.

Feijten, Peteke and Maarten Van Ham. 2007. "Residential mobility and migration of the divorced and separated." *Demographic Research* 17:623-654.

Ferguson, Mark R. and Pavlos S. Kanaroglou. 1995. "Utility Variability within Aggregate Spatial Units and its Relevance to Discrete Models of Destination Choice." Pp. 243-269 in *New Directions in Spatial Econometrics*, edited by L. Anselin and R. J. G. M. Florax. New York: Springer.

Fischer, Mary J. and Douglas S. Massey. 2004. "The Ecology of Racial Discrimination." *City and Community* 3:221-241.

Galle, Omer R. and Karl E. Taeuber. 1966. "Metropolitan Migration and Intervening Opportunities." *American Sociological Review* 31:5-13.

Greenwood, Michael J, Peter R Mueser, David A Plane, and Alan M Schlottmann. 1991. "New directions in migration research." *The Annals of Regional Science* 25:237-270.
Haynes, Kingsley E. and A. Stewart Fotheringham. 1984. *Gravity and Spatial Interaction Models*, Edited by G. I. Thrall. Beverly Hills: Sage.
Herting, Jerald R., David B. Grusky, and Stephen E. Van Rompaey. 1997. "The Social Geography of Interstate Mobility and Persistence." *American Sociological Review* 62:267-287.
Hipp, John R. 2009. "Specifying the Determinants of Neighborhood Satisfaction: A Robust Assessment in 24 Metropolitan Areas over Four Time Points." *Social Forces* 88:395-424.
—. 2010. "What is the “neighbourhood” in neighbourhood satisfaction? Comparing the Effects of Structural Characteristics Measured at the Micro-neighborhood and Tract Level." *Urban Studies* 47:2517-2536.
Hunt, Jennifer. 2004. "Are migrants more skilled than non-migrants? Repeat, return, and same-employer migrants." *Canadian Journal of Economics* 37:830-849.
Ihrke, David K. and Carol S. Faber. 2012. "Geographical mobility: 2005 to 2010." US Census Bureau, Washington, DC.
Jun, Il Soo and Hui S Chang. 1986. "Functional forms and the relevance of contiguous migration in the study of migration and employment growth." *The Annals of regional science* 20:17-27.
Kipnis, Baruch A. and Yoel Mansfeld. 1986. "Work-place utilies and commuting patterns: Are they class or place differentiated? " *The Professional Geographer* 38:160-169.
La Gory, Mark and John Pipkin. 1981. *Urban Social Space*. Belmont, CA: Wadsworth.
Lee, Barrett A., Karen E. Campbell, and Oscar Miller. 1991. "Racial Differences in Urban Neighboring." *Sociological Forum* 6:525-550.
Long, Larry. 1988. *Migration and Residential Mobility in the United States*. New York: Russell Sage.
Long, Larry, C Jack Tucker, and William L Urton. 1988a. "Measuring migration distances: Self-reporting and indirect methods." *Journal of the American Statistical Association* 83:674-678.
—. 1988b. "Migration distances: An international comparison." *Demography* 25:633-640.
Lu, Max. 1999. "Determinants of Residential Satisfaction: Ordered Logit vs. Regression Models." *Growth and Change* 30:264-287.
Massey, Douglas S. and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard.
McCarthy, J. F., M. Valenstein, and F. C. Blow. 2007. "Residential mobility among patients in the VA health system: associations with psychiatric morbidity, geographic accessibility, and continuity of care." *Adm Policy Ment Health* 34:448-55.
Mesch, Gustavo S. and Orit Manor. 1998. "Social Ties, Environmental Perception, and Local Attachment." *Environment and Behavior* 30:504-519.
Morrill, Richard L. and Forrest R. Pitts. 1967. "Marriage, Migration, and the Mean Information Field: A Study in Uniqueness and Generality." *Annals of the Association of American Geographers* 57:401-422.
Niedomysl, Thomas and Urban Fransson. 2014. "On Distance and the Spatial Dimension in the Definition of Internal Migration." *Annals of the Association of American Geographers* 104:357-372.

Mobility distance function

Niedomysl, Thomas, Ulf Ernstsson, and Urban Fransson. 2015. The accuracy of migration distance measures. *Population, Space and Place*. DOI: 10.1002/psp.1971

Pandit, Kavita. 2000. "Regional Variation in Mobility Levels and Timing in the United States." *The Professional Geographer* 52:483-493.

Plane, David A and Jason R Jurjevich. 2009. "Ties that no longer bind? The patterns and repercussions of age-articulated migration." *The Professional Geographer* 61:4-20.

Plane, David A. and Gordon F. Mulligan. 1997. "Measuring Spatial Focusing in a Migration System." *Demography* 34:251-262.

Quigley, John M and Daniel H Weinberg. 1977. "Intra-urban residential mobility: a review and synthesis." *International Regional Science Review* 2:41-66.

Ravenstein, Ernest G. 1885. "The laws of migration." *Journal of the Statistical Society of London* 48:167-235.

Rogers, Andrei, James Raymer, and Frans Willekens. 2003. "Imposing Age and Spatial Structures on Inadequate Migration-Flow Datasets." *The Professional Geographer* 55:56-69.

Simmons, James W. 1968. "Changing residence in the city: a review of intraurban mobility." *Geographical Review*:622-651.

South, Scott J. and Kyle D. Crowder. 1997. "Escaping Distressed Neighborhoods: Individual, Community, and Metropolitan Influences." *American Journal of Sociology* 102:1040-1084.

—. 1998. "Leaving the 'Hood: Residential Mobility between Black, White, and Integrated Neighborhoods." *American Sociological Review* 63:17-26.

Stillwell, John and Michael Thomas. 2015. "How far do internal migrants really move? Demonstrating a new method for the estimation of intra-zonal distance." *Regional Studies, Regional Science* 3:28-47.

Sun, Shipeng and Steven M Manson. 2010. "An agent-based model of housing search and intraurban migration in the Twin Cities of Minnesota." International Environmental Modelling and Software Society.

Taylor, P. J. 1975. "Distance Decay in Spatial Interactions. Concepts and Techniques in Modern Geography 2." *Norwich: Geo Abstracts.*

Taylor, Peter J. 1971. "Distances within Shapes: An Introduction to a Family of Finite Frequency Distributions." *Geografiska Annaler. Series B, Human Geography* 53:40-53.

Thomas, Michael, John Stillwell, and Myles Gould. 2015. "Modelling multilevel variations in distance moved between origins and destinations in England and Wales." *Environment and Planning A abstract* 47:996-1014.

Viry, Gil. 2012. "Residential mobility and the spatial dispersion of personal networks: Effects on social support." *Social Networks* 34:59-72.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Zipf, George Kingsley. 1949. *Human Behavior and the Principle of Least Effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley.

Table 1. Summary statistics for distance of moves for various subsamples

	Distance			Percent moved to same tract (2 different estimates)	
	Median	Mean	Std. Dev.	Upper bound	Lower bound
Full sample	4.4	114.2	422.9	21.5%	16.1%
Poverty	3.3	90.8	374.8	23.4%	18.2%
White	4.9	128.3	445.6	18.0%	13.3%
Black	3.2	65.9	330.9	24.3%	18.6%
Latino	3.3	60.3	278.1	21.2%	17.3%
Owner	6.4	116.8	421.5	16.0%	11.3%
Renters	3.7	112.4	423.9	21.6%	16.8%
Bachelors degree or more	(a)	166.8	513.7	15.6%	11.1%
More than HS degree, but not bachelors	(a)	115.3	427.5	18.0%	13.5%
High school degree	(a)	90.4	370.9	20.6%	15.6%
Less than HS degree	(a)	61.8	294.1	25.4%	20.3%
With kids 5 and under (but not older)	4.8	(a)	83.3	(a)	(a)
With kids 6 to 18 years old	4.5	(a)	236.2	(a)	(a)
Have no kids	4.6	(a)	163.4	(a)	(a)

Note: (a) estimates not available due to disclosure constraints.

Mobility distance function

Table 2. Outcome is logged distance of most recent move

Black	-0.504	**
	-(21.49)	
Latino	-0.323	**
	-(11.75)	
Other race	-0.276	**
	-(3.36)	
Asian	-0.249	**
	-(6.22)	
Years of education	0.079	**
	(26.53)	
Household income (logged)	0.097	**
	(6.31)	
Owner	0.071	**
	(3.99)	
Age (X 100)	-1.666	**
	-(5.31)	
Age squared (X 100)	2.251	**
	(6.76)	
Married	0.301	**
	(13.60)	
Widowed	0.082	*
	(1.98)	
Divorced	-0.007	
	-(0.28)	
Only children less than 6 years of age	-0.149	**
	-(6.04)	
Presence of children 6 to 18 years of age	-0.074	**
	-(3.83)	
Year: 1991	0.053	**
	(2.92)	
Year: 1995	0.076	**
	(4.18)	
Intercept	0.961	**
	(13.52)	

Mobility distance function

*** p < .01(two-tail test), * p < .05 (two-tail test), † p < .05 (one-tail test). T-values in parentheses.*

Figure Labels:

Figure 1. Percent of moves by mileage bins

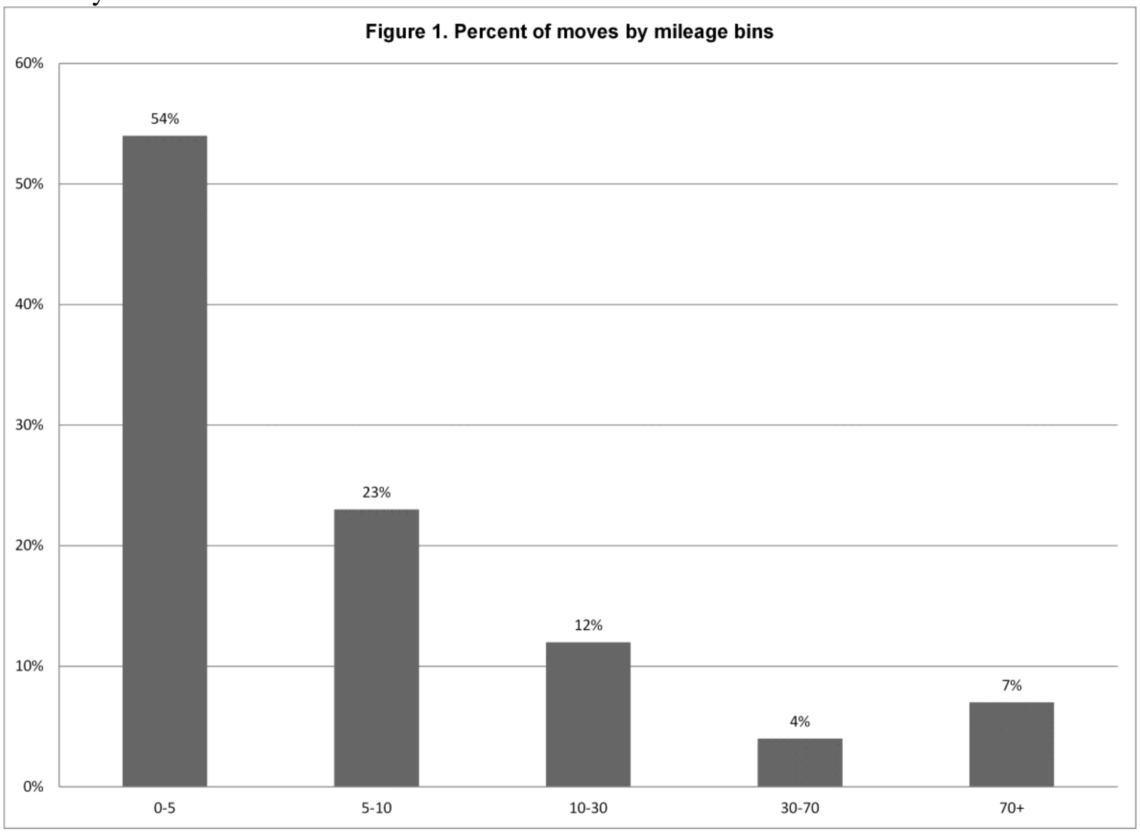
Figure 2. Logged distance of most recent move for full sample and those in poverty

Figure 3. Logged distance of most recent move for blacks, Latinos and whites

Figure 4. Logged distance of most recent move for households based on presence of children

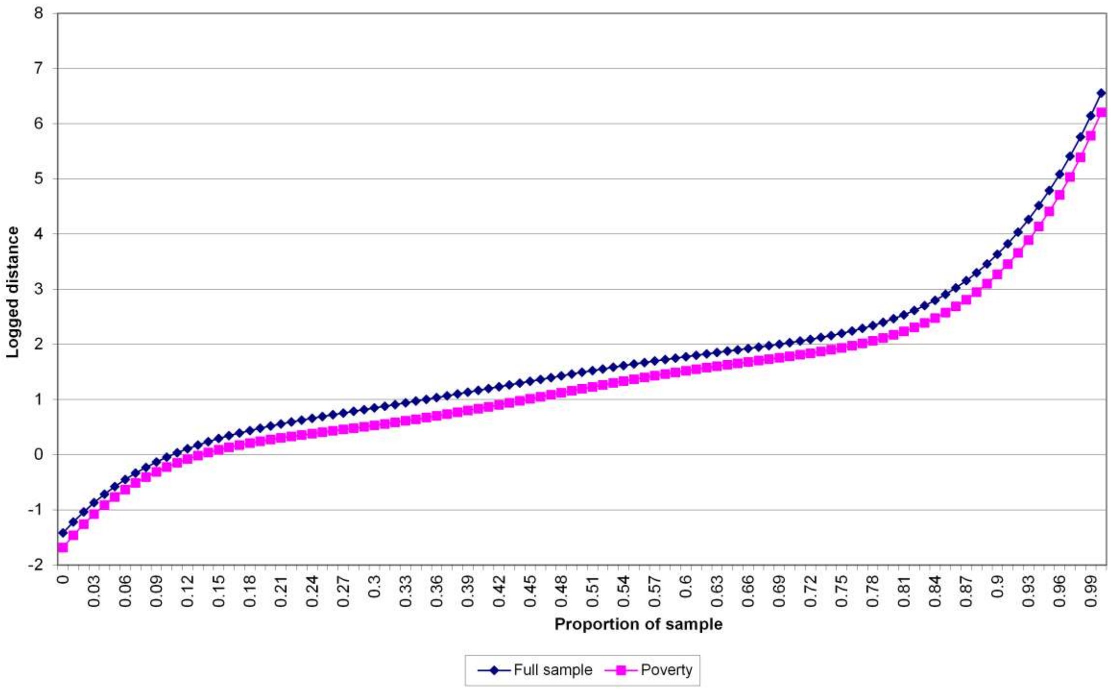
Figure 5. Logged distance of most recent move for owners and renters

Mobility distance function



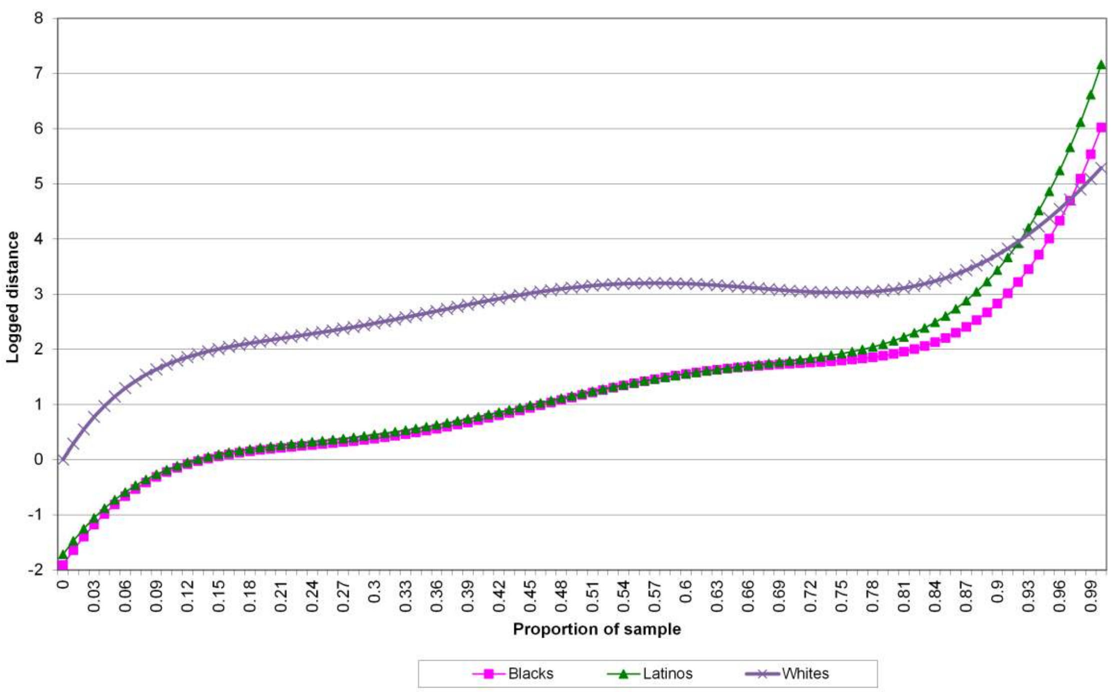
Mobility distance function

Figure 2. Logged distance of most recent move for full sample and those in poverty

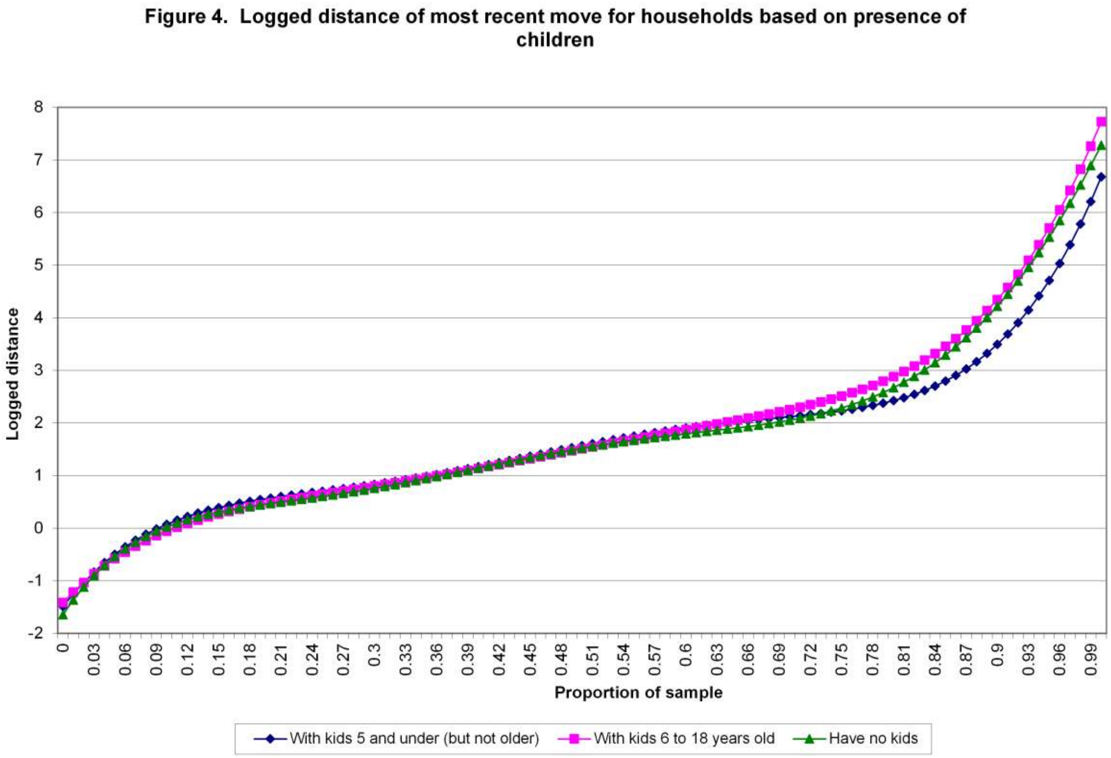


Mobility distance function

Figure 3. Logged distance of most recent move for blacks, Latinos and whites



Mobility distance function



Mobility distance function