

Residential Mobility and Location Choice

A Nested Logit Model with Sampling of Alternatives

Brian H. Y. Lee (Corresponding Author)
Assistant Professor
College of Engineering & Mathematical Sciences
University of Vermont
210 Colchester Avenue, Farrell Hall 114
Burlington, Vermont 05405-0303 U.S.A.
Phone: +1.802.656.1306
Fax: +1.802.656.9892
Email: bhylee@cems.uvm.edu

Paul Waddell
Professor
Department of City & Regional Planning
University of California, Berkeley
228 Wurster Hall #1850
Berkeley, California 94720-1850 U.S.A.
Email: waddell@berkeley.edu

TRB Paper #10-2533

Submission Date: 31 July 2009
Revision Submission Date: 15 November 2009

Word Count: 6969 in text + 2 tables (500) + 1 figures (250) = 7719

1 ABSTRACT

2 Residential mobility and relocation choice are essential parts of integrated transportation and land use
3 models. These decision processes have been examined and modeled individually to a great extent but
4 there remain gaps in the literature on the underlying behaviors that connect them. Households may partly
5 base their decision to move from or stay at a current location on the price and quality of the available
6 alternatives. Conversely, households that are on the market for a new location may evaluate housing
7 choices relative to their previous residence. How and the degree to which these decisions relate to each
8 other are, however, not completely understood. This research is intended to contribute to the body of
9 empirical evidence that will help answer these questions. It is hypothesized that residential mobility and
10 location choice are related household decisions that can be modeled together using a two-tier hierarchical
11 structure. This paper presents a novel nested logit (NL) model with sampling of alternatives and a
12 proposed procedure for sampling bias correction. The model was estimated using full-information
13 maximum likelihood estimation methods. The results confirm the applicability of this NL model and
14 support similar findings from other empirical studies in the residential mobility and location choice
15 literature.

17 INTRODUCTION

18 Transportation planning has traditionally focused on travel demand forecast models that assume land use
19 as exogenous inputs and ignore the feedback effects between changes in the transportation system and
20 urban development (1). Research advancements in this arena have been shifting away from these narrow
21 views towards more holistic approaches (2). The developments of large-scale integrated transportation
22 and land use microsimulation systems – such as ILUMASS (3, 4), ILUTE (5, 6), and UrbanSim (7, 8) –
23 have generated increasing attention on the research of a wide assortment of travel-related econometric
24 models. Two essential household behaviors that have been separately examined and modeled to a great
25 extent in recent years are residential mobility and location choice and they are the subjects of this
26 empirical study.

27 Residential mobility and location choice are critical components in systems of integrated
28 transportation and land use models. They are significant long-term household decisions that intimately
29 connect with the daily activity and travel aspects of individual lives (9, 10). Trade-offs between housing
30 qualities and rents, activity opportunities, and travel costs have long been recognized as fundamental
31 considerations in both the decision to move (11, 12) and the selection of a residence (13, 14).

32 In the transportation-land use modeling context, there are a few recent contributions on residential
33 mobility (e.g., 15, 16) and a considerable list of current research on residential location (e.g., 17-21).
34 Some studies have linked short- and medium-term travel-related behaviors with residential mobility (e.g.,
35 22-25) while numerous others made similar connections with residential location (e.g., 26-34). Few
36 studies have, however, explicitly modeled these closely related decisions together; one example used a
37 stated preference approach (35). None, to the best of the authors' knowledge, has modeled residential
38 mobility and location choice jointly in a hierarchical structure using revealed preference data, even
39 though these two parts of the housing decision process are most likely interdependent (36). Households
40 may partly base their decision to move from or stay at a current location on the price and quality of the
41 available alternatives (37). Conversely, households that are on the market for a new home may use their
42 most recent residence as a point of reference to help evaluate the location choices (20, 38). How and the
43 degree to which these decisions relate to each other are not completely understood. Further, there remain
44 opportunities for improvements in the modeling of this process. The state of the art integrated
45 microsimulation systems mentioned above all currently rely on separately and sequentially executed sub-
46 modules for simulating residential mobility and location choice; typically, the first sub-module uses
47 estimates of mobility rates to identify proportions of households to move and the latter sub-module
48 distributes them over space with models of location choice.

49 An important computational barrier that has limited the development of hierarchical residential
50 mobility and location models is the large number of potential alternatives commonly assumed to be
51 considered in the location decision. Depending on the granularity of the location choices, the universal set

of alternatives for a household looking to relocate can range from hundreds [e.g., neighborhoods or traffic analysis zones (19)] to hundreds of thousands [e.g., parcels or buildings (20)]. Such large numbers are computationally infeasible and behaviorally unrealistic. One practical solution to this intractability problem is to only use a sample of these alternatives for model estimation.

In a seminal work on residential location modeling, McFadden (39) demonstrated that the consistent estimation of choice models from a subset of alternatives is possible in the multinomial logit (MNL) form, where the property of independence from irrelevant alternatives (IIA) is assumed [see Ben-Akiva and Lerman (40, pp. 261-269) for a textbook treatment of this topic]. In more advanced, non-MNL model forms suitable for the analysis of multi-dimensional problems, the IIA restriction is relaxed to allow for correlations among alternatives. Consistent estimators of non-MNL models based on a sample of alternatives have not yet been proven in the literature (41). The latest research advances on this subject, however, are promising and suggest that a mathematical proof showing the consistent estimation of non-MNL choice models from a subset of alternatives may be possible (42-44).

This paper presents a two-tier nested logit (NL) model of residential mobility and location choice using household observations and building-level residences from the central Puget Sound region. It uses random sampling of alternatives and introduces a correction procedure to ensure that the nested logit maximum likelihood estimator provides consistent estimates of the parameters in the presence of random sampling, something that has not been done in prior research, to the authors' knowledge. This work intends to contribute to the body of empirical evidence on residential mobility and location behaviors and improve upon the understanding of the connection between them. It also offers a means for integrated transportation and land use microsimulation systems to jointly consider these parts in the housing process.

Following this introduction, sections detailing a review of the relevant literature on residential mobility and location choice are presented. They are then followed by a qualitative explanation of the NL formulation and the sampling correction procedure; a presentation of the data, model specification, and estimation results; and a discussion of the conclusions, limitations, and directions ahead.

REVIEW OF LITERATURE

The literature on residential mobility and location choice is extensive and spans across multiple related disciplines, with different expertise matched with different emphasis. Although these decisions are closely related in the household housing process, there is an expansive range of factors that contribute to each choice. Due to the overwhelming complexity of these behaviors and the assortment of circumstances, most empirical research are concerned with aspects of one decision or the other rather than with both components as a whole. In light of this division, these subjects are reviewed in the following separate, but overlapping, subsections.

Residential Mobility

Researchers have explored many different facets of residential mobility. For comprehensive reviews of this subject, readers are encouraged to survey compilation works by Clark (45), Clark and Dieleman (9), and Dieleman (10). Two of the most often cited studies in this literature are by Rossi (12) and Brown and Moore (11). The Rossi work is significant because it changed the focal point of residential mobility research from the study of aggregate spatial patterns to one of behaviors at the individual household level. This paradigm shift has directed most of the research performed since then and spawned a wealth of papers on the reasons for adjustments in household housing consumptions. One such work is by Brown and Moore, who decomposed residential mobility into two stages. The first concerns changes in household circumstances, both internal and external, that cause dissatisfaction with the existing housing choice. The second stage involves a decision to move or stay based on a survey of available alternatives in the housing stock. This conceptualization of the interdependence between mobility and location choice represents the process that the hierarchical model structure proposed in the current empirical study is attempting to capture.

In the decades since these important works, a multitude of studies on residential mobility were conducted in many different geographic, sociologic, and economic contexts (9, 46). Despite differences in

the environments, three intersecting findings transcend across geopolitical and social boundaries (10). The first is the close connection between mobility rates and the life cycle of an individual; young adults in their twenties and thirties compose the most mobile segment of the population (47, 48). The second concerns the relationship between mobility and both housing size and tenure; the rates are lower for households in larger homes and for those who own (9, 37). The third emphasizes the correlations between significant life course events and household moves; the acts of household formulation and dissolution and changes in education and work opportunities very often coincide with changes in residence (25, 46, 48).

Most recently, two contributions in the integrated transportation and land use modeling context are of note. M. Habib and Miller (16) introduced a conceptual process for residential mobility and spatial search that is to be implemented in the ILUTE microsimulation system (5, 6). In this paper, they focused on the mobility component and examined different discrete choice methods and hazard-based duration models. Their findings suggest that the use of different model formulations could reveal different reasons for household moves. Similarly, Eluru et al. (15) analyzed residential mobility behavior for motivations behind changes in housing consumption. In addition, they endogenized the length of time spent at a residence by jointly modeling reasons to move with the duration of stay. Their results support the existence of unobserved factors affecting both dimensions. These two works represent advances in disaggregate, agent-based modeling of residential mobility and contribute to the empirical understanding of this behavior. Nevertheless, they lack explicit feedback connections between mobility and location choice.

Residential Location Choice

The analysis of residential location choice at the household level was largely enabled by the development of discrete choice modeling methods. The early applications by Lerman (49) and McFadden (39) on this subject paved the way for a generation of research on identifying different contributing factors and making connections with travel-related behaviors. Much of this work is centered on the utility maximization concept where housing choice is represented as a bundle of other associated choices and households must make trade-offs between various housing qualities, neighborhood attributes, and accessibilities to best meet their needs.

Some studies concentrated on linking residential location with different modes and other travel choice behaviors (27, 28, 31, 32). Others have examined commuting factors and the relations between the locations of residence and workplace (22, 34). Abraham and Hunt (47), for one, have attempted to do both. The role of accessibility, to both work and non-work activities, in residential location choice has received a large share of the attention (17, 20, 26). Various modeling issues such as self-selection (29, 31) and price endogeneity (18) are also being addressed.

The multinomial logit (MNL) formulation mentioned in the introduction is only one of a family of discrete choice models but it has been, and still is, widely used in practice and research. The popularity of the MNL model in residential location choice applications is due mainly to properties that allow for efficient computations and consistent estimations with a subset of alternatives. Researchers have attempted to navigate around this sampling issue and use more advanced models with different strategies. K. Habib and Kockelman (50) estimated a series of nested logit (NL) and joint MNL models with full enumeration of all possible alternatives. Kim et al. (35) and Pinjari et al. (31) developed NL and joint binary-ordered logit models, respectively, by using categories of locations to significantly reduce the dimensionality. Others have simply used non-MNL models with sampling of alternatives and ignored sampling biases (51, 52); this strategy, as demonstrated by Nerella and Bhat (41) in their explorations of this subject, is highly discouraged.

The sampling issue is more pressing now than ever as research in integrated transportation and land use model systems require the use of disaggregate data and representations for increasing behavioral realism. Up until recent years, residential location has been commonly operationalized as aggregate units such as neighborhoods or traffic analysis zones (19). Although full enumerations of hundreds of these alternatives are undesirable, it remains, in some cases, computationally feasible. With the most recent location representations all the way down to the parcel or building levels (20, 21), however, the sampling

of alternatives for model estimation has become essential for practical use on current computers. It is this timely need that has helped motivate the direction for this research paper.

Nested Logit Model Formulation with Sampling of Alternatives

In this study, residential mobility and location choice are assumed to be two related stages of the household housing process that can be modeled jointly using a two-tier nested logit (NL) structure and full-information maximum likelihood estimation. Without any assumptions on the temporal sequence of the decisions, the nesting structure is assumed to include a binomial mobility choice of stay or move at the top level and a multinomial location choice at the bottom level. The stay nest only includes one location alternative (i.e., staying in the current location), whereas the move nest includes multiple choices for move locations. Figure 1 shows this residential mobility and location choice NL model.

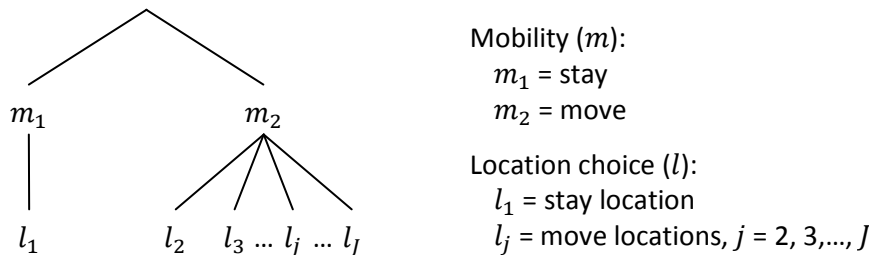


FIGURE 1 Two-tier nested structure of residential mobility and location choice.

The mathematical formulation of this model follows the utility maximizing NL model developed by McFadden (39), as described by Koppelman and Wen (53), but with an additional sampling correction procedure added to ensure consistent estimation of the model parameters using a subset of alternatives. The probability of a household choosing location l is defined as

$$P(l) = P(l|m) \cdot P(m), \quad (1)$$

where $P(l|m)$ is the conditional probability of l being chosen given nest m and $P(m)$ is the marginal probability of choosing nest m .

The bottom level conditional choice probability is equivalent to the standard multinomial logit (MNL) equation and has the form

$$P(l|m) = \frac{e^{V_l \mu_l}}{\sum_{l' \in L_m} e^{V_{l'} \mu_{l'}}}, \quad (2)$$

where V_l represents the observable components of the utility function for each elemental alternative l and μ_l is the associated scale parameter. When the simple random sampling strategy is used to draw a subset of alternatives, no correction is needed here as it has been shown that the standard MNL form produces consistent estimates of the model parameters due to the independence from irrelevant alternatives (IIA) property (39). For other sampling strategies, such as independent importance sampling or stratified importance sampling, different additional adjustment terms as described in Ben-Akiva and Lerman (40, pp. 261-269) are needed to correct for sampling bias.

The marginal choice probability of choosing nest m is

$$P(m) = \frac{e^{V'_m \mu_m}}{\sum_{m' \in M} e^{V'_{m'} \mu_{m'}}}, \quad (3)$$

where V'_m is the logsum (or inclusive value) associated with nest m and μ_m is the top level scale parameter. The logsum represents the expected value of the maximum of the random utilities of all alternatives in nest m . In a standard NL formulation with full enumeration of all available elemental alternatives, the logsum equals the log of the denominator of the conditional probability multiplied by $1/\mu_l$, or

$$V'_m = (1/\mu_l) \ln(\sum_{l' \in L_m} e^{V_{l'} \mu_{l'}}). \quad (4)$$

In this application, as is common in location choice models, the universal choice set is massive in size, and it is not computationally feasible to estimate the bottom level choice model using all available alternatives. If this was a MNL model of location choice only, then random sampling of alternatives provides a consistent estimate of the parameters of the choice model. The proposed model structure, however, is of the NL form with both residential mobility and location choice, and random sampling of alternatives is not possible without introducing sampling bias. In spite of this, prior research on residential location has employed NL with random sampling of alternatives, generally without any mention of the bias introduced by using sampling in a model structure that does not rely on the IIA property. This paper introduces a correction for this sampling bias by implementing a maximum likelihood estimator for the NL model that adjusts the logsum to account for the presence of sampling of alternatives at the bottom level of the nesting structure. By the Slutsky Theorem (40, p. 19), which states that a continuous function of consistent estimates is also consistent, an *expanded* logsum may be used as a consistent estimate of the logsum derived from the full set of alternatives. This is intuitive, since it simply means that it is necessary to account for the sample proportion within each nest in order to achieve a consistent estimate of the logsum for the nest. Failing to do so would provide a biased measure of the logsum. For simple random sampling, the expanded logsum is

$$V'_m = (1/\mu_l) \ln\{\sum_{l' \in L_m} [(1/R_{l'}) e^{V_{l'} \mu_{l'}}]\}, \quad (5)$$

where $R_{l'}$ is the sampling rate, $0 < R_{l'} < 1$, which only applies to the sampled non-chosen alternatives. The observed chosen alternative, by design, is not sampled and always included in the choice set and, thus, has an expansion factor of one. If other sampling strategies are used, then there may be different sampling rates for different sample groups and that should be reflected within the summation component of the expanded logsum calculations.

Note that of the two scale parameters, μ_l and μ_m , only one of them or the ratio between the two can be identified in the model estimation. It is necessary to normalize one of the parameters to equal one and estimate the other. An upper normalization where μ_m is set to one may be preferred for direct comparisons between NL and MNL models (54). Following the approach described in Koppelman and Wen (53), however, a lower normalization where μ_l is set to one is adopted here. In this case, μ_m must be between 0 and 1 as a condition of consistency. As the coefficient of the logsum, the parameter μ_m can be interpreted as an indicator of the hierarchical nature of the nesting structure. If the estimate of this parameter approaches 0, then the decision process is considered to be strictly hierarchical. Whereas if μ_m equals one, then the two choices are considered independent and the NL model reduces to a single-stage MNL model. In the case where there is only a single alternative in a nest (e.g., nest m_1 , the stay nest with only l_1 , the stay location), that nest is considered degenerate and μ_m equals one. If there is only a single alternative in every nest, then the NL model also collapses to the MNL model.

THE CENTRAL PUGET SOUND REGION EMPIRICAL EXAMPLE

The nested logit (NL) model of residential mobility and location choice described above was applied in the Seattle, Washington metropolitan area using a 2006 household activity survey. This application modeled intra-regional household movements and did not concern immigration from or emigration to locations outside of the study area. The model was estimated simultaneously using a new full-information maximum likelihood NL estimation package in the UrbanSim microsimulation system software (7, 8). The following subsections describe the data, explain the model specifications, and present the estimation results.

Data

The Puget Sound Regional Council (PSRC) 2006 Household Activity Survey (55) was the primary source of data for modeling residential mobility and location choice. This was complemented by 2005 building and parcel assessments from county tax assessors; 2000 business establishment data from the state

unemployment insurance database; 2000 census data; 2006 travel model results from the PSRC; and other regional geo-spatial data representing environmental, political, and planning features.

Executed as a two-day activity and travel diary, the PSRC 2006 survey obtained valid data from 4,739 households, representing 10,516 persons. The data includes x-y coordinate residential location information for all household residences at the time of the survey, which enabled assignment to individual buildings for modeling. The survey also contains the length of time each household has resided at that current home and, thus, it is possible to identify households who recently changed residences. Of the full set of surveyed households, a subset of 1,677 moved in the 5-year period prior to 2006 and the data has the x-y coordinate location of each of these households' previous home. This group of households who moved is considered to be the "Movers", while the rest of the surveyed households are the "Stayers".

The observed chosen alternative for both the Movers and the Stayers is the household residential location in 2006. In reference to the two-tier nested structure of the residential mobility and location choice model proposed in Figure 1, the chosen alternative for the Stayers is the location in the Stay nest (i.e., l_1 in nest m_1) and 29 other buildings were randomly sampled from the universal choice set to populated the move alternatives in the Move nest. For the Movers, there are two location data points: one is the 2006 location and the other is the previous residence. In this case, the 2006 location is considered l_2 in the Move nest and the previous residence is l_1 in the Stay nest. The remaining 28 alternatives (i.e., l_3, l_4, \dots, l_{30}) are randomly sampled.

In addition to the location information, there are other household and individual socio-demographical characteristics and activity and travel records that are relevant to the residential mobility and location choice model. This includes household income and size; the age and work/school status of each member; the nature of the residential tenure (i.e., rent or own); commuting times; and travel to non-work activities.

Model Specifications

Equations 1 through 3 and 5 from above were used for this Central Puget Sound application of the NL residential mobility and location choice model. The unit of analysis for the choice set is the individual residential building and the modeled choice behavior is that of the household. The attractiveness of location l for a household n is expressed by the utility function,

$$V_l = \alpha X_l + \beta(Y_n - P_l) + \gamma H_n X_l, \quad (6)$$

where α , β , and γ are parameters to be estimated; X_l is an array of characteristics describing alternative l ; Y_n is the household annual income; P_l is the housing price in annualized rents; and H_n is an array of attributes for household n . This utility expression highlights the underlying logic that when making a residential location decision, households take account of their own characteristics (e.g., household income and size, lifecycle stage) and those of the available housing options (e.g., price, accessibility). The interaction terms ($Y_n - P_l$ and $H_n X_l$) allow the household attributes (Y_n and H_n) to enter the choice model with the housing characteristics (P_l and X_l , respectively); the household attributes cannot be specified by themselves in the utility function because they do not vary across the alternatives and there would be no way to estimate coefficients for such variables. The income and price interaction ($Y_n - P_l$) represents household disposable income and is specified as such to allow housing price to enter into the model.

Based on a priori knowledge from the literature and previous work on a residential location choice model using the same data for this region (20), a wide assortment of explanatory variables were examined for their theoretical and statistical significance. The ones related to residential mobility include variables that were found to be some of the most important in the mobility literature (10), such as individual or household life stage (i.e., age), household size and composition, and residential tenure (i.e., rent or own). The variables relating to residential location choice that were explored include housing qualities (e.g., dwelling size per unit; building type such as single family residential or multi-family condominiums), neighborhood attributes (e.g., population density, proportions of household with children or young adults), and accessibility measures (e.g., general accessibility to regional work opportunities or local shopping activities, individual-level work travel time).

Model Estimation and Results

Full-information maximum likelihood techniques were used to estimate the parameters of the two-tier NL model of residential mobility and location choice. The log-likelihood function, with a lower normalization and an expanded logsum, is

$$\mathcal{L} = \sum_n \left\{ \sum_{m \in M} y_{mn} [V'_m \mu_m - \ln(\sum_{m' \in M} e^{V'_{m'} \mu_{m'}})] + \sum_{l \in L_m} y_{lmn} [V_l - \ln(\sum_{l' \in L_m} e^{V'_{l'}})] \right\}, \quad (7)$$

where y_{mn} is 1 if household n chooses nest m and 0 otherwise; y_{lmn} is 1 if household n chooses alternative l in nest m and 0 otherwise;

$$V'_m = \ln \left\{ \sum_{l' \in L_m} [(1/R_{l'}) e^{V'_{l'}}] \right\}; \quad (8)$$

and all else as defined previously.

After extensive experimentations with different specifications, one NL model was chosen based on the theoretical and statistical significance of the estimated parameters. The explanatory variables specified in this model and brief descriptions are summarized in Table 1 and the model estimates and results are shown in Table 2.

TABLE 1 Residential mobility and location choice model explanatory variables and descriptions

Variables	Descriptions
asc_stay	Alternative specific constant for the stay location (l_1 in nest m_1)
hh_avg_adult_age	Average age of all adults (18 years & older) in the household (HH)
hh_has_kids	Dummy of HH with kids (under 18 years)
hh_has_workers	Dummy of HH with full-time workers
hh_is_home_owner	Dummy of HH who are home owners
inc_less_price	Log of HH annual income less annual imputed rent per unit
kids \times SFR	Dummy of HH with kids \times dummy of single-family residential building
kids \times kids_HH	Dummy of HH with kids \times percent HH with kids within 600m
work_accessibility	Individual worker specific, zonal travel logsum for AM HBW drive alone trips to workplace location (maximum value between up to 2 workers)

TABLE 2 Model estimation results

Variables	Coefficient Estimates	T-values
<i>Stay alternative only:</i>		
asc_stay	-4.14	-23.3
hh_avg_adult_age	0.0589	8.29
hh_has_kids	0.0388	3.71
hh_has_workers	0.310	3.93
hh_is_home_owner	1.61	6.23
<i>Move alternatives only:</i>		
inc_less_price	0.00467	3.98
kids \times SFR	0.00802	2.67
kids \times kids_HH	0.00202	2.16
work_accessibility	0.0250	4.02
move nest logsum (μ_m)	0.0158	3.97
Observations (N)	4739	
Log-likelihood (\mathcal{L})	-7058.6	
Null log-likelihood	-47642.2	
LL ratio (ρ)	0.8518	
Adjusted ρ (ρ')	0.8516	

1 This NL model of residential mobility and location choice has a relatively parsimonious
2 specification but it, nevertheless, includes the important exploratory variables that are expected to be an
3 integral part of the household housing decision process. The estimated logsum parameter (μ_m) for the
4 Move nest is small at 0.0158, which strongly supports the hierarchical NL model structure. Trials of
5 models with fuller specifications generally resulted in μ_m estimates that were out of theoretical range.
6 This included experiments with specifying alternative specific variables for the Stay alternative that were
7 included in the Move nest. For this reason, only household specific variables were specified for the Stay
8 alternative.

9 The stay location alternative was considered the reference choice and four household specific
10 variables plus a constant were specified for this location. All of these variables have the expected positive
11 sign and they are all statistically significant. The first variable, hh_avg_adult_age, captures the life cycle
12 of the household based on the average age of the individual adult members. As expected, this estimated
13 variable parameter suggests that older households are more likely to stay than move, which corresponds
14 to the empirical evidence in the literature that says young adults are generally the most mobile segment of
15 the population (47, 48). The next two variables, hh_has_kids and hh_has_workers, are both indicators of
16 the household life cycle as well as the composition. The estimated parameters for these variables imply
17 that households with kids and workers are less likely to move than those without. These results echo other
18 studies that found having children and a steady job to be a constraint to the intention to move (9, 56). This
19 is likely due to inertial from a combination of social, professional, and educational ties associated with a
20 current residence. The last household specific variable is hh_is_home_owner and the estimated parameter
21 agrees with overwhelming empirical evidence from other studies, which points to home owners as being
22 less mobile than renters (10).

23 The next four variables are all interactions between household characteristics and location
24 attributes and they were specified for the move alternatives, using the stay location as a reference. The
25 first of these variables is inc_less_price, the household income and housing price interaction. Income is a
26 powerful indicator of a household's preference for and expected ability to afford different types of
27 housing. This variable, as described above in Equation 6 as $Y_n - P_l$, compares the household income to an
28 annual imputed rent and describes the amount of resources remaining for other consumption needs after
29 paying for housing. The price variable is theoretically vital in the decision making process but price
30 coefficients in location choice models have been reported as statistically insignificant, small, or even
31 positive in several empirical studies (18). In this and a previous analysis (20), the authors found this
32 specification of price with income to consistently provide statistically significant and theoretically sound
33 coefficients. The positive sign supports the logic that the more disposable income is associated with a
34 housing choice, the more attractive it is. The next two variables, kids \times SFR and kids \times kids_HH, express
35 the desires of households with children for single-family residential buildings and neighborhoods with
36 other children, respectively. These variables reflect the different needs and preferences between
37 households with kids and those without and relate to life course studies that correlate significant life
38 events such as having children with household changes in residence (25, 48, 57). The last variable is a
39 measure of accessibility that captures the ease of access between each residential location choice and
40 specific work locations, for each worker with a fixed non-home workplace. Since this measure is at the
41 individual-level while the residential mobility and location choice model is at the household-level, it was
42 aggregated by choosing the maximum value among the workers in multi-worker households. The positive
43 sign of the estimated parameter suggests that in general, good access to work is a positive contributing
44 factor both to the residential mobility and location choice decisions.

45 CONCLUSIONS

46 Residential mobility and location choice are interrelated decisions that make up the household housing
47 process. In this paper, they were modeled jointly using a two-tier nested logit (NL) structure with simple
48 random sampling of location alternatives and full-information maximum likelihood estimation. A
49 correction procedure to account for sampling bias was proposed and executed. To the best of the authors'
50 knowledge, this is the first empirical model of its kind, with the sampling correction, in the literature.
51

The model results from this research align with those found in previous empirical studies. Life stage, household composition, and housing tenure were determined to be important household characteristics that helped to explain the residential mobility and location choice process. In general, older households, ones with children or workers, and those who own their homes are less likely to move and change residences than younger renters or those with no kids or no workers. In addition, the housing price, the building type, the neighborhood composition, and accessibility were also found to be significant factors in the mobility-location choice calculus. Relative to a household's current home, location alternatives for a move are attractive if the cost is lower and the ease of access to work is higher, and, for households with children, if the building is a single-family residential home in a neighborhood with other families with children.

This NL model explicitly accounts for the multi-dimensional nature of the housing process and represents a step forward in the modeling of residential mobility and location choice in the integrated transportation and land use model system context. The interdependence between these decisions is reflected by the location alternatives in the Move nest informing the move-stay choice and the stay location being used as a reference to compare possible move locations. There are, however, some limitations to this model and they correspond to opportunities for future research in this arena. First, the housing tenure is considered as an exogenous input in this model, but tenure choice may be another interrelated part of the housing process that is better determined jointly with residential mobility and location choice. Second, the use of a cross-sectional household survey dataset with a retrospective view at housing allowed for the incorporation of previous residences in the model but it is no substitute for a panel dataset. Life course events and changes in individual and household circumstances were not observed and these factors are known to be important in the mobility literature when considering the household level of satisfaction with an existing housing situation (48, 57). Third, it should be noted that the household survey was conducted during a real estate price bubble and the behaviors observed may not reflect those in the current economic downturn. The consideration of macroeconomic conditions would help any predictive models in forecasting future behaviors. Lastly, the sampling of alternatives with corrections in a NL structure is a useful methodological enhancement but the model would likely improve with more sophisticated sampling procedures, advanced location choice set screening methods, and experimentations with different sample sizes.

ACKNOWLEDGEMENTS

Financial assistance was generously provided by the National Science Foundation (NSF EIA-0121326 and NSF EIA-0090832); the United States Environmental Protection Agency (EPA R831837); the United States Department of Transportation, Federal Highways Administration (FHWA Exploratory Advanced Research Program BAA DTFH61-07-R-00117). The authors wish to acknowledge the feedback of Professors Moshe Ben-Akiva, Joan Walker, and David Layton on the bias adjustment for random sampling of alternatives in a nested logit model estimator. We thank Liming Wang for his assistance in coding the model and data processing for this paper. Finally, we want to acknowledge the assistance of Hana Sevcikova in implementing the adjustment of the logsum in the Open Platform for Urban Simulation.

REFERENCES

1. Waddell, P., Ulfarsson, G. F., Franklin, J. P., & Lobb, J. (2007). Incorporating land use in metropolitan transportation planning. *Transportation Research Part A: Policy and Practice*, 41(5), 382-410.
2. Miller, E. J., Kriger, D. S., & Hunt, J. D. (1998). Integrated urban models for simulation of transit and land use policies. Washington, DC: Transit Cooperative Research Program, Transportation Research Board. Retrieved from <http://www.nap.edu/catalog/9435.html>.
3. Strauch, D., Moeckel, R., Wegener, M., Gräfe, J., Mühlhans, H., Rindsfuser, G., et al. (2005). Linking transport and land use planning: The microscopic dynamic simulation model ILUMASS. In P. M.

- 1 Atkinson, G. M. Foody, S. E. Darby, & F. Wu (Eds.), *Geodynamics* (pp. 295-311). Boca Raton,
2 Florida: CRC Press.
- 3 4. Wagner, P., & Wegener, M. (2007). Urban land use, transport, and environment models: Experiences
4 with an integrated microscopic approach. *disP*, 170, 45-56.
- 5 5. Miller, E. J., & Salvini, P. A. (2001). The integrated land use, transportation environment (ILUTE)
6 microsimulation modelling system: Description and current status. In D. A. Hensher (Ed.), *Travel*
7 *Behaviour Research: The leading edge* (pp. 711-724). Amsterdam; New York: Pergamon Press.
- 8 6. Salvini, P. A., & Miller, E. J. (2005). ILUTE: An operational prototype of a comprehensive
9 microsimulation model of urban systems. *Networks and Spatial Economics*, 5(2), 217-234.
- 10 7. Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation, and
11 environmental planning. *Journal of American Planning Association*, 68, 297-314.
- 12 8. Waddell, P. A., Borning, A., Noth, M., Freier, N., Becke, M., & Ulfarsson, G. (2003).
13 Microsimulation of urban development and location choices: Design and implementation of
14 UrbanSim. *Networks and Spatial Economics*, 3(1), 43-67.
- 15 9. Clark, W. A. V., & Dieleman, F. M. (1996). *Households and housing: Choice and outcomes in the*
16 *housing market*. New Brunswick, New Jersey: The Center for Urban Policy Research.
- 17 10. Dieleman, F. M. (2001). Modelling residential mobility; a review of recent trends in research. *Journal*
18 *of Housing and the Built Environment*, 16(3), 249-265.
- 19 11. Brown, L. A., & Moore, E. G. (1970). The intra-urban migration process: A perspective. *Geografiska*
20 *Annaler. Series B, Human Geography*, 52(1), 1-13.
- 21 12. Rossi, P. H. (1955). *Why families move: A study in the social psychology of urban residential*
22 *mobility*. Glencoe, Illinois: Free Press.
- 23 13. Alonso, W. (1964). *Location and land use: Toward a general theory of land rent*. Cambridge: Harvard
24 University Press.
- 25 14. Muth, R. F. (1969). *Cities and housing: The spatial pattern of urban residential land use*. Chicago:
26 University of Chicago Press.
- 27 15. Eluru, N., Sener, I. N., Bhat, C. R., Pendyala, R. M., & Axhausen, K. W. (2009). Understanding
28 residential mobility: A joint model of the reason for residential relocation and stay duration. In
29 *Transportation Research Board 88th Annual Meeting*. Washington, DC.
- 30 16. Habib, M. A., & Miller, E. J. (2008). Modeling residential mobility & spatial search behavior -
31 Estimation of continuous-time hazard & discrete-time panel logit models for residential mobility. In
32 *Transportation Research Board 87th Annual Meeting*. Washington, DC.
- 33 17. Chen, J., Chen, C., & Timmermans, H. (2008). Accessibility Trade-Offs in Household Residential
34 Location Decisions. *Transportation Research Record: Journal of the Transportation Research Board*,
35 2077, 71-79.
- 36 18. Guevara, C. A., & Ben-Akiva, M. E. (2006). Endogeneity in residential location choice models.
37 *Transportation Research Record: Journal of the Transportation Research Board*, 1977, 60-66.
- 38 19. Guo, J. Y., & Bhat, C. R. (2007). Operationalizing the concept of neighborhood: Application to
39 residential location choice analysis. *Journal of Transport Geography*, 15(1), 31-45.
- 40 20. Lee, B. H. Y., Waddell, P. A., Wang, L., & Pendyala, R. M. (forthcoming). Operationalizing time-
41 space prism accessibility in a building-level residential choice model: Empirical results from the
42 Puget Sound region. *Environment and Planning A*.
- 43 21. Waddell, P. (forthcoming). Modeling residential location in UrbanSim. In J. Preston, F. Pagliara, &
44 D. Simmonds (Eds.), *Modelling Residential Location Choice*. Hampshire, U.K.: Ashgate.
- 45 22. Clark, W. A. V., & Withers, S. D. (1999). Changing jobs and changing houses: Mobility outcomes of
46 employment transitions. *Journal of Regional Science*, 39(4), 653-673.
- 47 23. Clark, W. A. V., Huang, Y., & Withers, S. D. (2003). Does commuting distance matter? Commuting
48 tolerance and residential change. *Regional Science and Urban Economics*, 33(2), 199-221.
- 49 24. Prillwitz, J., Harms, S., & Lanzendorf, M. (2006). Impact of life-course events on car ownership.
50 *Transportation Research Record: Journal of the Transportation Research Board*, 1985, 71-77.

- 1 25. Prillwitz, J., Harms, S., & Lanzendorf, M. (2007). Interactions between residential relocations, life
2 course events, & daily commute distances. *Transportation Research Record: Journal of the*
3 *Transportation Research Board*, 2021, 64-69.
- 4 26. Ben-Akiva, M. E., & Bowman, J. L. (1998). Integration of an activity-based model system and a
5 residential location model. *Urban Studies*, 35(7), 1131-1153.
- 6 27. Eliasson, J., & Mattsson, L. (2000). A model for integrated analysis of household location and travel
7 choices. *Transportation Research Part A - Policy and Practice*, 34(5), 375-394.
- 8 28. Krizek, K. J. (2006). Lifestyles, residential location decisions, and pedestrian and transit activity.
9 *Transportation Research Record: Journal of the Transportation Research Board*, 1981, 171-178.
- 10 29. Krizek, K. J. (2003). Residential relocation and changes in urban travel: Does neighborhood-scale
11 urban form matter? *Journal of the American Planning Association*, 69(3), 265-281.
- 12 30. Ng, C. F. (2008). Commuting distances in a household location choice model with amenities. *Journal*
13 *of Urban Economics*, 63(1), 116-129.
- 14 31. Pinjari, A. R., Eluru, N., Bhat, C., Pendyala, R., & Spissu, E. (2008). Joint model of choice of
15 residential neighborhood and bicycle ownership: Accounting for self-selection and unobserved
16 heterogeneity. *Transportation Research Record: Journal of the Transportation Research Board*, 2082,
17 17-26.
- 18 32. Pinjari, A. R., Pendyala, R. M., Bhat, C. R., & Waddell, P. A. (2008). Modeling the choice
19 continuum: Integrated model of residential location, automobile ownership, bicycle ownership, &
20 commute tour mode choice decisions. In *Transportation Research Board 87th Annual Meeting*.
21 Washington, DC.
- 22 33. Sermons, M. W., & Koppelman, F. S. (2001). Representing the differences between female & male
23 commute behavior in residential location choice models. *Journal of Transport Geography*, 9, 101-110.
- 24 34. Waddell, P., Bhat, C., Eluru, N., Wang, L., & Pendyala, R. M. (2008). Modeling interdependence in
25 household residence & workplace choice. *Transportation Research Record: Journal of the*
26 *Transportation Research Board*, 2003, 84-92.
- 27 35. Kim, J. H., Pagliara, F., & Preston, J. (2005). The intention to move and residential location choice
28 behaviour. *Urban Studies*, 42(9), 1621-1636.
- 29 36. Wong, G. K. M. (2002). A conceptual model of the household's housing decision-making process:
30 The economic perspective. *Review Of Urban & Regional Development Studies*, 14(3), 217-234.
- 31 37. Clark, W. A. V., & Onaka, J. L. (1983). Life-cycle and housing adjustment as explanations of
32 residential-mobility. *Urban Studies*, 20(1), 47-57.
- 33 38. Zondag, B., & Pieters, M. (2005). Influence of accessibility on residential location choice.
34 *Transportation Research Record: Journal of the Transportation Research Board*, 1902, 63-70.
- 35 39. McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F.
36 Snickars, & J. Weibull (Eds.), *Spatial interaction theory and planning models* (pp. 75-96).
37 Amsterdam: North Holland.
- 38 40. Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel*
39 *demand*. MIT Press Series in Transportation Studies. Cambridge, Massachusetts; London, England:
40 The MIT Press.
- 41 41. Nerella, S., & Bhat, C. (2004). Numerical analysis of effect of sampling of slternatives in discrete
42 choice models. *Transportation Research Record: Journal of the Transportation Research Board*, 1894,
43 11-19.
- 44 42. Bierlaire, M., Bolduc, D., & McFadden, D. (2008). The estimation of generalized extreme value
45 models from choice-based samples. *Transportation Research Part B: Methodological*, 42(4), 381-394.
- 46 43. Frejinger, E., Bierlaire, M., & Ben-Akiva, M. (2009). Sampling of alternatives for route choice
47 modeling. *Transportation Research Part B: Methodological*, In Press.
- 48 44. Koppelman, F., & Garrow, L. (2005). Efficiently estimating nested logit models with choice-based
49 samples example applications: Example applications. *Transportation Research Record: Journal of the*
50 *Transportation Research Board*, 1921, 63-69.

- 1 45. Clark, W. A. V. (1982). Recent research on migration and mobility - A review and interpretation.
2 Progress in Planning, 18, 5-56.
- 3 46. Li, S., & Wu, F. (2004). Contextualizing residential mobility and housing choice: Evidence from
4 urban China. Environment and Planning A, 36(1), 1-6.
- 5 47. Abraham, J. E., & Hunt, J. D. (1997). Specification and estimation of nested logit model of home,
6 workplaces, and commuter mode choices by multiple-worker households. Transportation Research
7 Record: Journal of the Transportation Research Board, 1606, 17-24.
- 8 48. Clark, W. A. V., & Huang, Y. (2003). The life course and residential mobility in British housing
9 markets. Environment and Planning A, 35(2), 323 – 339.
- 10 49. Lerman, S. R. (1976). Location, housing, automobile ownership, and mode to work: A joint choice
11 model. Transportation Research Record: Journal of the Transportation Research Board, 610, 6-11.
- 12 50. Habib, K. M. N., & Kockelman, K. (2008). Modeling choice of residential location and home type:
13 Recent movers in Austin, Texas . In 87th Annual Meeting of the Transportation Research Board.
14 Washington, DC.
- 15 51. Yagi, S., & Mohammadian, A. (2008). Joint models of home-based tour mode and destination
16 choices: Applications to a developing country. Transportation Research Record: Journal of the
17 Transportation Research Board, 2076, 29-40.
- 18 52. Zhou, B., & Kockelman, K. (2008). Microsimulation of residential land development and household
19 location choices: Bidding for land in Austin, Texas. Transportation Research Record: Journal of the
20 Transportation Research Board, 2077, 106-112.
- 21 53. Koppelman, F. S., & Wen, C. (1998). Alternative nested logit models: Structure, properties and
22 estimation. Transportation Research Part B, 32(5), 289-298.
- 23 54. Carrasco, J. A., & Ortuzar, J. D. (2002). Review and assessment of the nested logit model. Transport
24 Reviews, 22(2), 197-218.
- 25 55. Cambridge Systematics Inc. (2007). Final report: PSRC 2006 household activity survey analysis
26 report. Sammamish Wash.: Washington Department of Transportation.
- 27 56. Kim, J., Pagliara, F., & Preston, J. (2003). An analysis of residential location choice behaviour in
28 Oxfordshire, UK-a combined state preference approach. International Review of Public
29 Administration, 8(1), 103-114.
- 30 57. Li, S. (2004). Life course and residential mobility in Beijing, China. Environment and Planning A,
31 36(1), 27 – 43.