# MODELING RESIDENTIAL SORTING EFFECTS TO UNDERSTAND THE IMPACT OF THE BUILT ENVIRONMENT ON COMMUTE MODE CHOICE

Abdul Rawoof Pinjari
Department of Civil, Architectural & Environmental Engineering
The University of Texas at Austin
1 University Station, C1761
Austin, Texas 78712
Phone: (512) 964-3228; Fax: (512) 475-8744
Email: abdul.pinjari@mail.utexas.edu

Ram M. Pendyala, Ph.D.
Department of Civil & Environmental Engineering
Arizona State University
PO Box 875306, ECG252
Tempe, AZ 85287-5306
Phone: (480) 727-9164; Fax: (480) 965-0557
Email: ram.pendyala@asu.edu

Chandra R. Bhat, Ph.D.
Department of Civil, Architectural & Environmental Engineering
The University of Texas at Austin
1 University Station, C1761
Austin, Texas 78712
Phone: (512) 471-4535; Fax: (512) 475-8744
Email: bhat@mail.utexas.edu

&

Paul A. Waddell, Ph.D.
Center for Urban Simulation and Policy Analysis
Daniel J. Evans School of Public Affairs
University of Washington
Box 353055
Seattle, Washington 98195-3055
Phone: (206) 221-4161; Fax: (206) 685-9044
Email: pwaddell@u.washington.edu

**ABSTRACT**

This paper presents an examination of the significance of residential sorting or self selection effects in understanding the impacts of the built environment on travel choices. Land use and transportation system attributes are often treated as exogenous variables in models of travel behavior. Such models ignore the potential self selection processes that may be at play wherein households and individuals choose to locate in areas or built environments that are consistent with their lifestyle and transportation preferences, attitudes, and values. In this paper, a simultaneous model of residential location choice and commute mode choice that accounts for both observed and unobserved taste variations that may contribute to residential self selection is estimated on a survey sample extracted from the 2000 San Francisco Bay Area household travel survey. Model results show that both observed and unobserved residential self selection effects do exist; however, even after accounting for these effects, it is found that built environment attributes can indeed significantly impact commute mode choice behavior. The paper concludes with a discussion of the implications of the model findings for policy planning.

**Keywords:** causality, heterogeneity, joint model, built environment, residential self-selection, travel behavior

## 1. INTRODUCTION

The importance and the complexity of the land use - travel behavior relationship has been recognized for several decades in the transportation planning practice and research communities. The complexity of the land use - travel behavior association arises due to (1) the multitude of dimensions that define land use (for example, land use mix, urban form, street block density, and local network features) and travel behavior (such as auto ownership, mode choice, and overall travel demand), and (2) the possibility of multiple causal and/or pure associative relationships between the dimensions that define land use and travel behavior (see Bhat and Guo, 2007 for an extended discussion on the land use – travel behavior relationship).

In conventional transportation planning practice, a one-way causal flow in which the nature of the land use pattern affects travel behavior is often assumed. Assuming such a one-way causal relationship would mean that households and individuals first locate themselves in neighborhoods based on market forces such as housing affordability, crime statistics, and school quality. Their travel behavior is then shaped by neighborhood characteristics (or built environment attributes). The above reasoning would imply, for example, that land use patterns and neighborhood attributes can be modified to achieve a desired shift in travel mode shares. The fallacy in such a one-way cause-and-effect assumption, which implies a sequential nature of residential location and mode choice decisions (in that order), is that it ignores the associative nature of the decisions. That is, the relationship between residential location and travel mode choice decisions may be a mix of partial cause-and-effect linkage and partial associative correlation. In reality, households and individuals may locate themselves into neighborhoods that allow them to pursue their activities using modes that are compatible with their socio-demographics (*e.g.*, income), attitudes (*e.g.*, auto-disinclination), and travel preferences (*e.g.*, preference for smaller commute time). If this is indeed the case, then urban land-use policies aimed at modifying neighborhood attributes for inducing mode shifts would alter the spatial residential location patterns more than the mode choice patterns. This phenomenon is called

2

residential self selection or residential sorting and calls for the treatment of residential location choice as an endogenous choice dimension that needs to be modeled simultaneously with the travel behavior dimension of interest. Ignoring the endogeneity of residential location choice or residential sorting effects (when present), can result in the identification of "spurious" causal effects of neighborhood attributes on travel behavior and lead to distorted policy implications. In order to correctly assess the impact of land-use patterns on mode choice, one must recognize and control for the associative correlations that may arise due to residential sorting. In light of this discussion, the specific objectives of this study are to:

- Clearly understand the mechanism of the relationship between residential location patterns and commute mode choice.

- Assess the impact of built environment (BE) attributes on mode choice by controlling for residential sorting effects and disentangling the "spurious" and "true" causal effects of the neighborhood attributes on commute mode choice.

In order to accomplish the objectives, a comprehensive analysis of the effect of neighborhood attributes on commute mode choice is undertaken through a joint residential location choice and mode choice modeling effort. An extensive suite of neighborhood attributes or descriptors are used for the analysis of built environment effects as are a range of demographic variables in the mode choice model. In addition, a key aspect of the modeling framework employed in this paper is that both observed and unobserved heterogeneity (*i.e.*, sensitivity variations due to household/individual observed demographics and unobserved factors) are accommodated in analyzing the effect of neighborhood attributes on residential location choice and mode choice.

The econometric modeling methodology used in this paper is an extension of the general joint modeling methodology developed recently by Bhat and Guo (2007), in which they control for the endogeneity of residential location patterns (*i.e.*, self selection effects) to assess

3

the impact of neighborhood attributes on car ownership. In that paper, car ownership is treated as an ordered discrete response choice variable. The modeling framework proposed in this paper is different in that the travel behavior variable of interest here (mode choice) is of an unordered discrete response nature.

The contribution of this paper is thus two-fold. First, the joint model can control for residential sorting effects to obtain the "true" effect of neighborhood attributes on mode choice. Such a joint model can predict the spatial residential relocation patterns as well as the travel behavior (mode choice in this case) changes that may be brought about in response to land-use policies. Second, from a methodological standpoint, the paper presents a methodology for simultaneously modeling the relationship between two unordered multinomial discrete choice variables, thus accommodating both causal as well as associative components of the relationship that may exist between them (residential location choice and commute mode choice in the current context). This is the first self-selection study that the authors are aware of in which two unordered discrete choice variables are modeled using a joint analysis framework.

The remainder of the paper is organized as follows. Following a brief review of the literature in the next section, the modeling methodology is presented in the third section. In the fourth section, a description of the data used in the study is presented. Model results are presented in the fifth section together with a discussion of the interpretation of the findings. Finally conclusions are presented in the sixth and final section.


## 2. LITERATURE REVIEW

There is a vast body of literature dedicated to the relationship between land use and travel behavior (for a review of the literature, see Ewing and Cervero, 2001, Bhat and Guo, 2007, Transportation Research Board – Institute of Medicine, 2006, and Cao and Mokhtarian, 2006). This section highlights some of the previous work germane to the topic addressed in this paper, *i.e.*, the relationship between residential location choice and mode choice.

4

Numerous studies in the past have examined the impact of neighborhood attributes on mode choice.  Several of them (for example, see Friedman *et al.*, 1994, Frank and Pivo, 1994, Ewing *et al.*, 1994, Handy, 1996, Cervero and Wu, 1997, Cervero and Kockelman, 1997, Kockelman, 1997, Badoe and Miller, 2000, Crane, 2000, Ewing and Cervero, 2001, Rajamani *et al.*, 2003, and Rodriguez and Joo 2004, and Zhang, 2004) reported a significant impact of neighborhood attributes in mode choice decisions. However, not all earlier studies have found such significant impacts of neighborhood attributes. For instance, Crane and Crepeau (1998) and Hess (2001) found no evidence that land use affects travel mode choice patterns. Kitamura *et al.* (1997) examined the effects of land use, demographic, and attitudinal variables on the proportion and number of trips by various modes, and found that attitudinal and demographic variables dominate neighborhood attributes in their effects on travel mode choice.  Cervero (2002) studied mode choice behavior in Montgomery County, Maryland and found that the influences of urban design tend to be more modest than those of intensities and mixtures of land use on mode choice decisions.

Most of the studies listed above ignore residential sorting effects when estimating the impact of neighborhood characteristics on travel mode choice. However, there are a few exceptions. Boarnet and Sarmiento (1998), for example, accounted for residential sorting effects through an instrumental variable technique in their analysis of non-work auto trip making.  Their findings, using data from southern California region, indicate a rather weak impact of built environment effects on non-work travel by auto mode, after accounting for residential self-selection. Cervero and Duncan (2002) accommodated for residential self-selection by estimating a nested logit model for the joint choices of residing near a rail station and commuting by rail transit. Their analysis with the 2000 San Francisco Bay Area data suggests that residential sorting due to transit-oriented lifestyle preferences accounts for about 40 percent of the rail-commute decision. Cervero and Duncan (2003), in another study accounting for residential self-selection in the San Francisco Bay area, found that the impact of neighborhood

attributes diminishes considerably after accounting for residential sorting effects. Zhang (2006) accommodated for residential sorting effects through an instrumental variable approach in his joint model of auto ownership, residential location, and travel mode choice. His analysis indicates that auto dependency is highly sensitive to street network connectivity and automobile availability. Schwanen and Mokhtarian (2005) found that, though residential sorting plays a significant role in explaining commute mode choice, neighborhood characteristics have a non-negligible effect on commute mode choice even after controlling for such self selection effects.

In the context of residential self selection, the recent work by Bhat and Guo (2007) offers a comprehensive and general methodology to control for residential sorting effects. Specifically, they control for residential sorting due to observed socio-demographic and unobserved factors in an ordered response model of household car ownership (See Bhat and Guo, 2007 for an explanation of the advantages of this methodology over other methods of accommodating residential self-selection). The current study builds upon Bhat and Guo's work by developing a joint model of residential location choice and mode choice that explicitly accommodates residential sorting effects and accounts for both observed and unobserved heterogeneity in residential self-selection. A detailed explanation of the methodology follows in the next section.

## 3. ECONOMETRIC MODELING FRAMEWORK

### 3.1 Mathematical Formulation

The equation system for the joint residential location choice and commute mode choice model may be written as follows:

$$u^*_{hi} = \gamma'_h x_i + \varepsilon_{hi}, \text{ spatial unit } i \text{ chosen if } u^*_{hi} > \max_{\substack{k=1,2,...I \\ k \neq i}} u^*_{hk} \tag{1}$$

$$\mu^*_{q_h rj} = \alpha'_{q_h j} y_{q_h} + \beta'_{q_h} z_{q_h rj} + \delta'_{hj} x_r + \xi_{q_h rj}, \text{ mode } j \text{ chosen if } \mu^*_{q_h j} > \max_{\substack{m=1,2,...J \\ m \neq j}} \mu^*_{q_h m}$$

The utility expressions in the equation system (1) can be rewritten as the following equation system (the reader is referred to Table 1 for a quick reference of the terms used in Equations 1 and 2):

$$u_{hi}^{*} = \sum_{l}\left(\gamma_{l} + \Lambda_{l}^{'}w_{hl} + v_{hl}\right)x_{il} + \left(\sum_{l}\omega_{hl}x_{il} + \varepsilon_{hi}\right) \tag{2}$$

$$\mu_{q_{h}rj}^{*} = \alpha_{q_{h}j}^{'}y_{q_{h}} + \beta_{q_{h}}^{'}z_{q_{h}rj} + \sum_{l}\left(\delta_{jl} + \Delta_{jl}^{'}s_{hl} + \eta_{hjl}\right)x_{rl} + \left(\sum_{l}\pm\omega_{hjl}x_{rl} + \zeta_{q_{h}rj}\right)$$

**Table 1 about here**

The first equation in the equation systems (1) and (2) is the utility function for the choice of residence in which $u_{hi}^{*}$ is the indirect utility that the household $h$ derives from locating itself in spatial unit $i$, $x_i$ is a vector of attributes corresponding to spatial unit $i$ ($x_i$ can potentially include non-built environment (non-BE) attributes such as racial composition, commute time, *etc.* and built environment (BE) attributes such as land-use mix, density, transit-accessibility, *etc.*), and $\gamma_h$ in equation system (1) is a household-specific coefficient vector capturing the sensitivity to attributes in vector $x_i$. $\gamma_h$ is parameterized in the first equation of the equation system (2) as: $\gamma_{hl} = (\gamma_l + \Lambda_l^{'}w_{hl} + v_{hl} + \omega_{hl})$, where $w_{hl}$ is a vector of observed household-specific factors affecting sensitivity to the $l^{th}$ attribute in vector $x_i$, and $v_{hl}$ and $\omega_{hl}$ are household-specific unobserved factors impacting the sensitivity of household $h$ to the $l^{th}$ attribute. $v_{hl}$ includes only those household-specific unobserved factors that influence sensitivity to residential choice, while $\omega_{hl}$ includes only those household-specific unobserved factors that impact both residential choice and commute mode choice. Finally, $\varepsilon_{hi}$ is an idiosyncratic error term assumed to be identically and independently extreme-value distributed across spatial alternatives $i$ and households $h$.

The second equation in equation systems (1) and (2) is the utility function for the choice of commute mode in which $\mu^{*}_{q_h rj}$ is the indirect utility that an individual $q$ from household $h$ residing in spatial unit $r$ associates with commute mode $j$. In the explanatory variables, $y_{q_h}$ is a vector of attributes that includes non-spatial determinants of modal utilities such as individual and household level socio-demographics (for example, household and personal income, age, gender, *etc.*), $z_{q_h rj}$ is a vector of level-of-service (LOS) attributes faced by the individual $q$ of household $h$ between his/her observed residential location $r$ and employment location by mode $j$ (for example, travel time, travel cost, *etc.*), and $x_r$ is a vector of attributes corresponding to the chosen residential spatial unit $r$ (for example, BE attributes such as land-use mix, density, *etc.*, and household level non-BE attributes such as the total commute time of all commuters in the household).

In the coefficient vectors in the second equation of the equation systems (1) and (2), $\alpha_{q_h j}$ represents the impact of socio-demographics on the utility of mode $j$, $\beta_{q_h}$ is a vector of response sensitivities to the LOS attributes in $z_{rq_h j}$, and $\delta_{hj}$ is a household-specific coefficient vector capturing the impact of BE and non-BE attributes (in vector $x_r$) of chosen residential spatial unit $r$ on the utility of mode $j$. The elements (indexed by $l$) of $\delta_{hj}$ are parameterized in the second equation of the equation system (2) as: $\delta_{hjl} = (\delta_{jl} + \Delta'_{jl}s_{hl} + \eta_{hjl})$, where $s_{hl}$ is a vector of observed household-specific factors influencing the sensitivity to $l^{th}$ attribute in $x_r$, $\Delta_{jl}$ is the corresponding vector of coefficients, and $\eta_{hjl}$ is a term capturing the impact of household-specific unobserved factors on the sensitivity to $l^{th}$ attribute in $x_r$. Finally, $\xi_{q_h j}$ of the equation system (1) is an error term that is partitioned into two components in the equation system (2) as: $\sum_{l}(\pm\omega_{hjl})x_{rl} + \zeta_{q_h j}$. The $\pm\omega_{hjl}x_{rl}$ terms are the common error components in

residential choice and mode choice, while $\zeta_{q_h j}$ is an idiosyncratic term assumed to be identically and independently (IID) logistic distributed across individuals and modal alternatives.

**3.2 Intuitive Discussion of Model Structure**

In the equation system (2), the self-selection of households into certain neighborhoods (that explains the endogeneity in the effect of neighborhood specific BE and non-BE attributes on commute mode choice) is captured by controlling for both observed and unobserved factors that impact residential location and commute mode choice. The explanation is as follows.

First, the model formulation controls for the effect of systematic/observed socio-demographic differences among individuals in their mode choice decisions. Suppose households with high income avoid residing in high density neighborhoods. This can be reflected by including income as a variable in the $w_{hl}$ vector in the residential choice equation. High income households are also likely to own more cars and the individuals belonging to those households are more likely to choose auto as their commute mode choice. The residential sorting based on income can then be controlled for when evaluating the effect of the BE attribute "density" on commute mode choice by including income as a variable in the $y_{q_h}$ vector in the mode choice equation. Ignoring such residential sorting effects due to observed demographics can lead to an artificial inflation of the neighborhood attribute effects in mode choice decisions.

Second, the model formulation controls for unobserved attributes (such as attitudes/perceptions, and environmental considerations) that may influence both residential choice and commute mode choice. For example, households with individuals that are environment-conscious and auto-disinclined may locate themselves into neighborhoods that are conducive to the use of non-motorized forms of transport so that they may walk or bike to work. Such common unobserved preferences are captured in the terms $\omega_{hl}$ and $\omega_{hjl}$ of the residential

9

choice utility equations and the non-motorized modal utility equations, respectively. These common unobserved factors cause the endogeneity in the effect of corresponding BE and non-BE attributes in the commute mode choice model, and give rise to correlation in the error components across the residential location and mode choice models leading to the joint nature of the model structure.

The '$\pm$' in front of the $\omega_{hjl}x_{rl}$ terms in the mode choice equation indicates that the impact of common unobserved factors in moderating the influence of the characteristics represented by $x_{rl}$ across the residential choice and mode choice equations may be in the same or opposite directions, respectively (called as positive or negative correlation, respectively). If the sign is '+', it implies that the unobserved factors that increase (decrease) the individuals' (households) preference to the characteristic represented by $x_{rl}$ in residential location choice decisions also increase (decrease) their preference for commute mode $j$, while a '−' sign implies that the unobserved factors that increase (decrease) the individuals' preference to the characteristic captured by $x_{rl}$ in residential location choice decisions decrease (increase) their preference for commute mode $j$.

If the $x_{rl}$ measures are defined in the context of promoting smart growth and neo-urbanism concepts (such as high density and increased land use diversity) to promote non-motorized travel to work, then there may be an expectation that the appropriate sign in front of the $\omega_{hjl}x_{rl}$ term in non-motorized modal utility equations should be positive. Through the model formulation adopted in this paper, it is possible to test which one of the two signs is appropriate. A positive sign suggests that households who have an intrinsic preference for neo-urbanist neighborhoods also have a higher preference for non-motorized modes of transport (due to unobserved attributes such as auto-disinclination). Ignoring these $\omega_{hjl}x_{rl}$ terms while estimating the mode choice utility equations leads to an artificial inflation of the positive sign on the

corresponding neo-urbanist BE attributes (*i.e.*, an artificial inflation of the positive sign on the $\delta_{jl}$ terms in the non-motorized modal utility equations).

If $x_{rl}$ represents an attribute such as total commute time of all individuals in the household, the anticipated sign in front of the $\omega_{hjl} x_{rl}$ term in auto modal utility equations could be either positive or negative. A negative sign indicates that the unobserved factors (such as attitudes/perceptions towards traveling and spending time on the road) that increase (decrease) individuals' sensitivity to total commute time in residential location decisions also increase (decrease) their preference for the relatively faster auto modes. On the other hand, a positive sign indicates the presence of unobserved factors affecting residential location choice that contribute to individuals/households increasing their total commute time and therefore becoming more auto-oriented in their commute mode choice. For example, one may consider such factors as crime, school quality, aesthetic appeal of neighborhood, neighborhood amenities, and perceptions of the prestige associated with living in a certain neighborhood. Although individuals/households would like to minimize their total commute time index, simply doing so may result in their locating in less-desirable residential neighborhoods. These unobserved factors then lead to individuals/households living in neighborhoods that increase their total commute time index and make them more auto-oriented.

In summary, the model formulation explicitly considers residential sorting effects that may be traced to observed socio-demographics, and unobserved attitudinal variables and personal lifestyle preferences. An important note on causality and the joint nature of residential location and mode choice decisions is in order here. As it can be seen from the modal utility part of the Equation 2, the characteristics of the "chosen" residential location are being used in the commute mode choice model. That is, the commute mode choice is modeled conditional upon the residential location decisions. This implies a hierarchy that residential location decisions precede commute mode choice decisions. Thus, the model structure assumes a causal

influence of the residential location choice (and hence the built environment) on commute mode choice. Along with this hierarchy (or the causal structure), households and individuals may locate (or self-select) themselves in built environments (or residential locations) that are consistent with their socio-demographics, lifestyle preferences, attitudes and values. This self-selection phenomenon leads to endogeneity representative of a behaviorally joint decision process. Self-selection (and hence the behaviorally joint decision process) may occur either due to observed factors such as socio-demographics, or due to unobserved factors such as attitudes and values. Thus, by including observed and unobserved factors that affect both residential choice and mode choice decisions, the residential self-selection phenomenon (and hence the behaviorally joint nature of the decision process) is accounted for. Within the context of unobserved factors, the presence of common unobserved factors leads to an econometrically joint model structure. In other words, the model structure assumes that the residential location choice and mode choice decisions are made jointly, but with an in-built hierarchy that the residential location choice affects mode choice. Considering the long-term nature of the residential location choice decisions, it is reasonable to assume a hierarchy (*i.e.,* a causal structure) that residential location choice affects commute mode choice.

### 3.3 Model Estimation

The parameters to be estimated in the equation system (2) include the $\alpha$ and $\beta$ vectors, the $\gamma_l, \delta_l, \Lambda_l$, and $\Delta_l$ vectors, and the variances of $v_{hl} (= \sigma_{vl}^2), \eta_{hjl} (= \sigma_{\eta l}^2)$, and $\omega_{hl} (= \sigma_{\omega l}^2)$ for those BE and non-BE attributes with random taste heterogeneity. In a general case, where $\sigma_{vl}^2 \neq 0$, $\sigma_{\eta l}^2 \neq 0$, and $\sigma_{\omega l}^2 \neq 0$ for each of the BE and non-BE attributes (i.e., for each $l$), there may be unobserved factors that affect the sensitivity to each of the BE and non-BE attributes, which are specific to residential location choice, mode choice, as well as common to both residential location and mode choices. However, in specific empirical cases, it is to be noted that the

random taste heterogeneity to a particular attribute $l$ may occur only in residential choice ($\sigma_{vl}^2 \neq 0$, $\sigma_{\eta l}^2 = 0$, $\sigma_{\omega l}^2 = 0$), only in some of the modal utilities ($\sigma_{vl}^2 = 0$, $\sigma_{\eta l}^2 \neq 0$, $\sigma_{\omega l}^2 = 0$), independently in residential choice and mode choice ($\sigma_{vl}^2 \neq 0$, $\sigma_{\eta l}^2 \neq 0$, $\sigma_{\omega l}^2 = 0$), or as combinations of the above patterns with a common effect on both residential choice and mode choice ($\sigma_{\omega l}^2 \neq 0$). Also, there may not be any random heterogeneity for some or all of the attributes in either of the residential choice and mode choice models ($\sigma_{vl}^2 = 0$, $\sigma_{\eta l}^2 = 0$, $\sigma_{\omega l}^2 = 0$).

Let $\Omega$ represent a vector that includes all the parameters to be estimated, and let $\Omega_{-\sigma}$ represent a vector of all parameters except the variance terms. Also, let $c_h$ be a vector that stacks the $v_{hl}$, $\eta_{hjl}$, and $\omega_{hl}$ terms across all BE and non-BE attributes and let $\Sigma$ be a corresponding vector of standard errors. Define $a_{hi} = 1$ if household $h$ resides in spatial unit $i$ and 0 otherwise. Similarly, define $b_{q_h j} = 1$ if an individual $q_h$ chooses the commute mode $j$ and 0 otherwise. Then, the likelihood function for a given value of $\Omega_{-\sigma}$ and $c_h$ may be written for an individual $q_h$ as:

$$L_{q_h}(\Omega_{-\sigma}) \mid c_h = \left[\frac{\exp(\gamma_h' x_i)}{\sum_k \exp(\gamma_h' x_k)}\right]^{a_{hi}} \left[\frac{\exp(\alpha_{q_h j}' y_{q_h} + \beta_{q_h}' z_{q_h rj} + \delta_{hj}' x_r)}{\sum_k \exp(\alpha_{q_h j}' y_{q_h} + \beta_{q_h}' z_{q_h rj} + \delta_{hj}' x_r)}\right]^{b_{q_h j}} \tag{3}$$

Finally, the unconditional likelihood function can be computed for individual $q_h$ as:

$$L_{q_h}(\Omega) = \int_{c_h} \left(L_{q_h}(\Omega_{-\sigma}) \mid c_h\right) dF(c_h \mid \Sigma), \tag{4}$$

where $F$ is the multidimensional cumulative normal distribution. The log-likelihood function can be written as: $L(\Omega) = \sum_{q_h} \ln L_{q_h}(\Omega)$. Simulation techniques are applied to approximate the multidimensional integral in Equation (4), and maximize the resulting simulated log-likelihood function. Specifically, the scrambled Halton sequence (see Bhat, 2003) is used to draw

realizations of $c_h$ from its population normal distribution. In the current paper, 125 realizations of $c_h$ were used to obtain stable estimation results.

## 4. DATA

### 4.1 Data Sources

The primary data source used in the analysis is the 2000 San Francisco Bay Area Travel Survey (BATS), designed and administered by MORPACE International, Inc. for the Bay Area Metropolitan Transportation Commission (see MORPACE International Inc., 2002 for details on survey design, sampling, and administration procedures). In addition to the activity survey, six other data sets associated with the San Francisco Bay area were used in the current analysis: land-use/demographic coverage data, zone-to-zone network level-of-service (LOS) data, a GIS layer of bicycle facilities, the Census 2000 Tiger files, census demographic data, and Public Use Microdata Sample (PUMS) data. Bhat and Guo (2007) offer a detailed explanation of the various data sources and how they were used to construct an integrated and comprehensive land use – travel behavior – LOS database that can be used to study land use – travel behavior relationships. The following section provides a description of the estimation sample.

### 4.2 Estimation Sample

The geographic area of study in this research is the Alameda County in the San Francisco Bay Area with 233 transport analysis zones. The residential choice of households and commute mode choice of individuals within this county constitute the focus of analysis for this paper. After extracting the Alameda County households from the survey sample and merging the various secondary data sources, the final sample for analysis comprised 1,878 individuals from 1,447 households.

This sample of 1,878 individuals includes only commuters who are employed outside the home. The average age of the sample persons is 43 years and about 56 percent of the persons are male. More than 85 percent of the individuals are employed full time. A vast majority (97.9%) is licensed to drive. The mode shares in the sample are as follows: a majority of the commuters (82.1%) drive alone, about 11 percent carpool either as a driver (4.7%) or passenger (6%), less than one percent (0.7%) use transit, and about 6.5 percent use non-motorized modes (2.8% bike and 3.8% walk) to commute to and from work.

The 1,878 individuals belong to 1,447 households with an average household size of about 2.5 persons per household, and with nearly a quarter of the households reporting household sizes of four or more persons. About one-third of the households report having an individual less than 18 years of age in the household. The median household income is rather high with about 50 percent of the households falling into the fourth and highest income quartile. On average, households reported a little over two cars per household with less than two percent of the households having zero cars. On average, the ratio of vehicles to licensed drivers is greater than one, generally indicating a high level of auto availability. A little less than two-thirds of the households own bicycles while about one-quarter of the households have three or more bicycles.

## 5. MODEL ESTIMATION RESULTS

This section provides a description of the model estimation results. The model system is estimated as a joint choice model including both residential location choice and commute mode choice dimensions. All 233 zones are considered to be alternatives in the residential location choice set. The commute mode choice set definition accounts for modal availability at the individual/household level. A household must own an automobile and an individual must have a driver's license for the auto drive (drive alone and drive with passenger) modes to be available in the choice set. The auto-passenger mode choice is available to all individuals as are the bike

and walk modes. The transit mode is included in the choice set based on transit availability (between residential and work zones) as specified in the network level of service files.

Table 2 presents estimation results for the residential location choice model. In general, the results are found to be plausible and consistent with expectations. The first variable in Table 2, logarithm of the number of households in a zone is a surrogate measure for the number of housing opportunities in a zone. As expected, a positive coefficient on this variable indicates that households are more likely to locate in zones with larger number of housing opportunities. Similarly, households are more likely to locate in zones with high household density. However, it is found that seniors are less likely to locate in zones of high density as evidenced by the negative coefficient associated with the interaction term. As expected high employment density zones are less likely to be chosen for residential location, except for lower income households who may be compelled to choose lower cost housing in such locations. Also, households desiring to live in single family detached housing units are more likely to locate in zones with a higher fraction of such a housing stock. The land use mix measure is negatively associated with residential location choice; this suggests that households are more prone to live in zones that are rather homogeneous in nature. This finding may also be an artifact of both zoning policies and zone definition strategies. Zoning policies may often dictate that land uses be segregated and traffic analysis zones themselves are often defined based on homogeneity of land uses. As a result, the likelihood of a household being located in a mixed land use zone is potentially going to be small simply because such zones are few and far between. Rather surprisingly (but consistent with the findings in Bhat and Guo, 2007), the fraction of residential land area is negatively associated with residential location choice. A higher recreational accessibility is associated with a greater likelihood of locating residence in a particular zone.

**Table 2 about here**

The total drive commute time for the household serves as a surrogate measure of the overall location of the household *vis-à-vis* the work locations of the commuters in the household

(assuming work locations are exogenous). Thus, this variable may be treated as an overall commute time index for the household. As expected, households attempt to locate such that this commute time index is reduced as evidenced by the negative coefficient associated with this variable. The total drive commute cost variable is found to be significant for households in the lowest quartile suggesting that lower income households are more sensitive to commuting costs than other households.

Within the context of the commute time index, the standard deviation of its random coefficient specific to the residential location model is highly significant with a test statistic value of 11.82, indicating significant population heterogeneity in the sensitivity to commute time index in residential location decisions. It is also found that there are common unobserved factors affecting both residential location choice and auto mode (all auto modes) choice in the context of commute time index; the corresponding error components are found to be negatively correlated. The standard error of this negative error correlation is found to be marginally significant with a test statistic value of 1.53. The presence of this correlation suggests that it is very important to model residential location choice and mode choice in a simultaneous equations framework because there are unobserved factors related to commute time that affect both of these choice dimensions simultaneously. In this particular instance, the interpretation of the negative sign on the correlation is as follows. The unobserved factors that increase (decrease) the sensitivity of individuals/households to total commute time index in residential location decisions, also make them more (less) oriented towards the relatively faster auto modes. For example, one may consider such factors as individuals' attitudes/perceptions towards traveling and spending time on the road that could contribute to higher (lower) sensitivity to total commute time index in residential location decisions, as well as higher (lower) preference to auto modes. Not accounting for such endogeneity could potentially lead to biased estimates of the impact of total commute time index in the commute mode choice model.

Within the context of common unobserved factors, only the total drive commute time variable has common random coefficients representing residential self-selection effects due to unobserved factors. It is possible that there may be important but omitted neighborhood variables (due to unavailability in the data) that might have resulted in significant unobserved residential self-selection effects associated with them. Further, an analysis in a different context may indicate the presence of unobserved residential self-selection effects (and hence an econometrically joint nature of the residential location and mode choice model) and/or random heterogeneity in sensitivity with respect to several neighborhood attributes. In any case, even with a comprehensive set of neighborhood attributes, it is important to estimate the joint model to test for the presence of unobserved residential sorting effects.

The remaining variables in Table 2 offer plausible interpretations consistent with expectations. Among the network level of service measures, street block density, bicycle facility density, availability of transit service to work zone, and the ease of access to a transit stop are desirable attributes with respect to residential location choice. However, as expected, households with higher vehicle availability are likely to be those located in suburban zones with lower street block density. This is supported by the negative coefficient associated with the interaction term between street block density and household vehicle availability. Similarly, the positive coefficient associated with the interaction term between bicycle facility density and bicycle ownership indicates that households with higher bicycle ownership are likely to be located in zones with higher bicycle facility density. Although transit availability is itself positively influencing residential location choice, transit stop access time negatively impacts residential location choice. This finding is not surprising in that while most zones are served by transit, most households are living in suburban locations where the access time to a stop is likely to be greater.

The demographic, housing cost, and ethnic composition variables all indicate that there is a natural self-selection process that occurs in the housing market. Similar income groups,

similar ethnic groups, and households of similar size tend to cluster together.  The median housing value has a negative impact on residential location choice suggesting that, as housing prices increase, the likelihood of locating in a zone decreases.

Results of the mode choice model estimation are presented in Table 3.  All of the results are plausible and consistent with expectations.  Relative to the auto mode, all other modes are less preferred as evidenced by the negative alternative specific constants.  Higher vehicle availability is associated with auto mode usage while higher bicycle ownership is positively associated with bicycle mode usage. Higher household sizes are associated with the use of shared-ride modes consistent with the greater opportunity and/or need for sharing a ride when there are multiple individuals in a household. Both travel time and travel cost have negative coefficients, with an added negative effect in the absence of work arrangement flexibility. Presumably, sensitivity to travel time becomes more pronounced in the absence of work flexibility.

**Table 3 about here**

The total drive commute time for the household serves as a surrogate for the location of the household *vis-à-vis* the work locations of the workers in the household.  The positive coefficient here is consistent with the notion that as households locate themselves such that their overall distance to the workplace increases, then the likelihood of becoming auto-oriented with respect to commute mode choice increases as well.  The standard error of the negative error correlation term in the context of the total drive commute time index variable is suggestive of the influence of common unobserved factors that affect residential location choice and choice of auto modes.  The interpretation and explanation of this finding was presented earlier in the context of the description of the results of Table 2.

Higher population and employment density contribute positively to bicycle and walk mode usage while a higher degree of land use mix contributes positively to transit usage. Similarly, a higher street block density and bicycle facility presence contribute positively to the

use of non-motorized modes of transportation. It is to be noted here that the current model specification allows for the process of households *self selecting* themselves into neighborhoods with street block density (and bicycle facility density) compatible with their vehicle availability (and bicycle ownership). The control for such residential sorting is achieved by including vehicle availability and bicycle ownership variables in the mode choice model. These findings are consistent with those in the literature and suggest that, even when controlling for residential sorting effects, the built environment attributes (street block density and bicycle facility presence in this case) have non-negligible effects on commute mode choice.

Log-likelihood ratio tests were performed to assess the significance and contribution of observed factors and unobserved residential sorting (joint correlation) effects. The log-likelihood value at convergence for the final joint model is -9384.7. The corresponding value for the model with no allowance for unobserved variations in sensitivity to the built environment and commute attributes is -9430.94. Then, the likelihood ratio test for testing the presence of unobserved variations in sensitivity is 92.47, which is larger than the critical chi-square value with 2 degrees of freedom at any reasonable level of significance (the 2 degrees of freedom correspond to the standard deviations on the drive commute time coefficient in the residential location model, and on the common error component, related to drive commute time coefficient, between the residential location and mode choice models). Further, the log-likelihood value corresponding to equal probability for each of the 233 zonal alternatives in the residential location model and sample shares in the car ownership model (corresponding to the presence of only the threshold parameters) is -11494.3. Therefore, the likelihood ratio index for testing the presence of exogenous variable effects and unobserved taste variations is 4219, which is substantially larger than the critical chi-square value with 38 degrees of freedom at any level of significance. Overall, these test results indicate that residential sorting effects are significant as are observed and unobserved taste variations in explaining commute mode choice behavior.

## 6. SUMMARY AND CONCLUSIONS

This paper addresses the key role of residential sorting effects in studying the impact of built environment attributes on travel mode choice. In the current land use – transportation planning context where the merits of altering the structure of the built environment to bring about changes in travel behavior are being debated, this study makes an important contribution to the field by presenting a joint model of residential location choice and commute mode choice that accounts for both observed and unobserved self-selection processes.

In previous studies of land use – travel behavior relationships, the residential location choice dimension is treated as exogenous and travel characteristics are often assumed to be affected by the attributes of the residential location. These studies often ignore the residential self-selection process that may be taking place in the housing market. Households/individuals may be locating in certain neighborhoods due to their lifestyle preferences, attitudes, values, and other unobserved factors. In the presence of such residential sorting effects, one may erroneously overestimate the impacts of built environment attributes on travel choices. In reality, individuals and households may simply be locating in neighborhoods that offer attributes consistent with their intrinsic preferences, attitudes, and values. More recent work in the field has recognized this important concept and begun to attempt to account for residential sorting effects in evaluating the impacts of the built environment on travel behavior.

This paper presents a rigorous econometric methodological framework for simultaneously modeling residential location choice and commute mode choice, two endogenous unordered multinomial discrete choice variables, while accounting for both observed and unobserved heterogeneity in the choice processes. The model system is estimated on a sample of households and individuals residing in Alameda County who responded to the activity-based household travel survey conducted in the San Francisco Bay Area in 2000.

The model estimation results offer some key conclusions that shed additional light on the debate surrounding the land use – travel behavior relationship. First, it is found that there are significant observed factors contributing to residential self selection. It is found that households self select their residential location based on demographic characteristics such as auto and bicycle ownership, income, household size, and race. Second, and more importantly, the common error component on the total drive commute time variable supports the endogenous treatment of residential location choice in a simultaneous equations modeling framework. The negative error correlation associated with this variable suggests that there are unobserved factors that may increase (decrease) the sensitivity of households and individuals to overall commute time in their residential location decisions and also make them more (less) auto-oriented in their commute mode choice decisions. Third, and perhaps most importantly, the built environment attributes such as accessibility, density, and land use mix have significant impacts on commute mode choice even after controlling for residential sorting effects and unobserved taste variations that contribute to such effects.

From a policy perspective, the results suggest that built environment attributes are not truly exogenous in travel choice decisions made by individuals. Households and individuals are locating themselves in built (transportation) environments that are consistent with their lifestyle preferences, attitudes, and values. In other words, households and individuals are making residential location and travel choice decisions jointly as part of an overall lifestyle package. Nevertheless, the findings in this paper suggest that modifying the built environment can bring about changes in mode choice behavior as evidenced by the significance of these attributes in the commute mode choice model even after controlling for residential sorting effects.

This research can be extended in at least three directions. First, it is important to carryout a subsequent policy simulation study to; (1) assess the extent of the impact of built environment policies, and (2) to assess the benefits accrued by accounting for residential sorting effects. Second, use of rich data sets with attitudinal variables may enhance the

22

understanding of the built environment – commute mode choice relationship. Third, the study relies upon statistical association between revealed choices as a means to assess the cause-and-effect relationship between the corresponding decisions. While such revealed choice data provides information on the observed decisions of decision-makers, it does not provide insights into the underlying behavioral processes that lead to those decisions (Ye *et al.*, 2007). In order to clearly understand the underlying behavior, detailed data on behavioral processes and decision sequences is needed.

**ACKNOWLEDGEMENTS**

**REFERENCES**

Badoe, D.A., Miller, E.J.: Transportation-Land-Use Interaction: Empirical Findings in North America, and Their Implications for Modeling. Transport. Res. D **5**(4), 235-263 (2000).

Bhat, C.R.: Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. Transport. Res. B **37**(9), 837-855 (2003).

Bhat, C.R., Guo, J.Y.: A Comprehensive Analysis of Built Environment Characteristics on Household Residential Choice and Auto Ownership levels. Transport. Res. B **41**(5), 506-526 (2007).

Boarnet, M.G., Sarmiento, S.: Can Land-use Policy Really Affect Travel Behavior? A Study of the Link between Non-work Travel and Land-Use Characteristics. Urban Studies **35**(7), 1155-1169 (1998).

Cao, X., Mokhtarian, P. L., Handy, S. L.: Examining the impacts of residential self-selection on travel behavior: Methodologies and empirical findings. Paper presented at the 11th International Association for Travel Behavior Research, Kyoto, August 2006.

Cervero, R.: Built Environments and Mode Choice: Toward a Normative Framework. Transport. Res. D **7**(4), 265-284 (2002).

Cervero R., Duncan, M.: Residential Self Selection and Rail Commuting: A Nested Logit Analysis. Working paper. University of California Transportation Center, Berkeley, CA, 2002. http://www.uctc.net/papers/604.pdf

Cervero, R., Duncan, M.: Walking, Bicycling, and Urban Landscapes: Evidence from the San Francisco Bay Area. Am. J. Public Health **93**(9), 1478-1483 (2003).

Cervero, R., Kockelman, K.: Travel Demand and the Three D's: Density, Diversity and Design. Transport. Res. D **2**(3),199-219 (1997).

Cervero, R., Wu, K.: Influences of Land Use Environments on Commuting Choices: An Analysis of Large U.S. Metropolitan Areas using the 1985 American Housing Survey. Working paper. University of California Transportation Center, Berkeley, CA, 1997. http://www.uctc.net/papers/669.pdf

Crane, R. The Influence of Urban Form on Travel: An Interpretive Review. J. Planning Literature **15**(1), 3-23 (2000).

Crane, R., Crepeau, R.: Does Neighborhood Design Influence Travel? A Behavioral Analysis of Travel Diary and GIS Data. Transport. Res. D **3**(4), 225-238 (1998).

Ewing, R., Cervero, R.: Travel and the Built Environment – Synthesis. Transport. Res. Rec. **1780**, 87-114 (2001).

Ewing, R., Haliyur, P., Page, W.: Getting Around a Traditional City, a Suburban Planned Unit Development, and Everything in Between. Transport. Res. Rec. **1466**, 53-62 (1994).

Frank, L. D., Pivo, G.: Impacts of Mixed Use and Density on the Utilization of Three Modes of Travel: Single Occupant Vehicle, Transit and Walking. Transport. Res. Rec. **1466**, 44-52 (1994).

Friedman, B., Gordon, P., Peers, J.: Effect of Neotraditional Neighborhood Design on Travel Characteristics. Transport. Res. Rec. **1466**, 63-70 (1994).

Handy, S.: Methodologies for Exploring the Link between Urban Form and Travel Behavior. Transport. Res. D **1**(2), 151-165 (1996).

Hess, D.: Effect of Free Parking on Commuter Mode Choice - Evidence from Travel Diary Data. Transport. Res. Rec. **1753**, 35-42 (2001).

Kitamura, R., Mokhtarian, P.L., Laidet, L.: A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. Transportation **24**(2),125-158 (1997).

Kockelman, K.M.: Travel Behavior as a Function of Accessibility, Land Use Mixing and Land Use Balance: Evidence from the San Francisco Bay Area. Transport. Res. Rec. **1607,** 116-125 (1997).

MORPACE International, Inc. Bay Area Travel Survey Final Report, March 2002. ftp://ftp.abag.ca.gov/pub/mtc/planning/BATS/BATS2000/

Rajamani, J., Bhat, C.R., Handy, S., Knaap, S., Song, Y.: Assessing Impact of Urban Form Measures on Nonwork Trip Mode Choice After Controlling for Demographic and Level-of-Service Effects. Transport. Res. Rec. **1831**, 158-165 (2003).

Rodriguez, D.A., Joo, J.: The Relationship between Non-motorized Mode choice and the Local Physical Environment. Transport. Res. D **9**(2), 151-173 (2004).

Schwanen, T., Mokhtarian, P.L.: What Affects Commute Mode Choice: Neighborhood Physical Structure or Preferences toward Neighborhoods? J. Transport Geog. **13**(1), 83-99 (2005).

Transportation Research Board and Institute of Medicine (TRB-IOM): Does the Built Environment Influence Physical Activity? Examining the Evidence. January, 2005. http://onlinepubs.trb.org/onlinepubs/sr/sr282.pdf.

Ye, X., Pendyala, R.M., Gottardi, G.: An Exploration of the Relationship Between Mode Choice and Complexity of Trip Chaining Patterns. Transport. Res. B **41**(1), 96-113 (2007).

Zhang, M.: The Role of Land Use in Travel Mode Choice: Evidence from Boston and Hong Kong. J. American Planning Assoc. **70**(3), 344-360 (2004).

Zhang, M.: Travel Choice with No Alternative: Can Land Use Reduce Automobile Dependence? J. Planning Education and Res. **25**(3), 311-326 (2006).

**TABLE 1. Description of Terms Used in Equations 1 and 2**

| | |
|---|---|
| $h$ | subscript for household $h$ |
| $q_h$ | subscript for individual $q$ from household $h$ |
| $i$ | subscript for any residential spatial unit $i$ |
| $r$ | subscript for the chosen residential spatial unit |
| $j$ | subscript for any mode $j$ |
| $l$ | subscript for $l^{th}$ attribute |
| $x_{il}$ | $l^{th}$ neighborhood attribute of spatial unit $i$, used in residential utility |
| $x_{rl}$ | $l^{th}$ neighborhood attribute of chosen spatial unit $r$, used in modal utility |
| $w_{hl}$ | vector of socio-demographic attributes affecting sensitivity to $l^{th}$ neighborhood attribute ($x_{il}$) in residential utility |
| $y_{q_h}$ | vector of socio-demographic attributes affecting modal utility |
| $z_{q_h rj}$ | vector of commute level-of-service (LOS) attributes by mode $j$ between the chosen residential and work locations |
| $s_{hl}$ | vector of socio-demographic attributes affecting sensitivity to $l^{th}$ neighborhood attribute ($x_{rl}$) in modal utility |
| $\gamma_l$ | sensitivity to $l^{th}$ neighborhood attribute ($x_{il}$) in residential utility |
| $\delta_{jl}$ | sensitivity to $l^{th}$ neighborhood attribute ($x_{rl}$) in modal utility |
| $\Lambda'_l$ | vector of coefficients on $w_{hl}$, indicating heterogeneous sensitivity to $l^{th}$ neighborhood attribute ($x_{il}$) in residential utility |
| $\Delta'_{jl}$ | vector of coefficients on $s_{hl}$, indicating heterogeneous sensitivity to $l^{th}$ neighborhood attribute ($x_{rl}$) in modal utility |
| $\alpha'_{q_h j}$ | vector of coefficients on socio-demographics ($y_{q_h}$) in modal utility |
| $\beta'_{q_h}$ | vector of coefficients on LOS attributes ($z_{q_h rj}$) in modal utility. This vector can be parameterized to capture heterogeneity. |
| $\nu_{hl}$ | mode specific error component capturing unobserved factors affecting the sensitivity to $l^{th}$ neighborhood attribute ($x_{il}$) |
| $\eta_{hjl}$ | error component capturing unobserved factors affecting the sensitivity to $l^{th}$ neighborhood attribute ($x_{rl}$) in residential utility |
| $\omega_{hl}$ | common error component capturing common unobserved factors affecting the sensitivity to $l^{th}$ neighborhood attribute |

**TABLE 2. Estimation Results of the Residential Location Choice Model**

| Variables | Parameter | t-stat |
|---|---|---|
| **Zonal size and density measures (including demographic interactions)** | | |
| Logarithm of number of households in zone ($\times 10^{-1}$) | 9.803 | 15.02 |
| Household density (#households per acre x $10^{-1}$) | 0.351 | 3.70 |
|     Interacted with presence of seniors in household | -0.652 | -1.93 |
| Employment density (#employment per acre x $10^{-1}$) | -0.211 | -2.89 |
|     Interacted with household income in the lowest quartile | 0.196 | 2.38 |
| **Zonal land-use structure variables (including demographic interactions)** | | |
| Fraction of residential land area | -0.813 | -5.70 |
| Fraction of single family housing interacted with household living in single family detached housing | 2.298 | 13.03 |
| Land-use mix | -0.305 | -2.07 |
| **Regional accessibility measures (including demographic interactions)** | | |
| Recreation accessibility x $10^{-2}$ (by auto mode) | 0.425 | 6.35 |
| **Commute-related variables  (including demographic interactions)** | | |
| Total drive commute time of all commuters in household (minutes x $10^{-2}$) | -11.472 | -24.28 |
|     Standard deviation of the error term in residential location model | 5.809 | 11.82 |
|     Standard deviation of the error term common to residential location and mode choice models (negative correlation between the error terms) | 0.859 | 1.53 |
| Total drive commute cost of all commuters in household (dollars x $10^{-1}$) | 0 | fixed |
|     Interacted with household income in the lowest quartile | -4.600 | -2.47 |
| **Local transportation network measures (including demographic interactions)** | | |
| Street block density (number of block per square mile x $10^{-2}$) | 0.163 | 1.47 |
|     Interacted with number of vehicles per number of licenses in household | -3.526 | -3.34 |
| Bicycle facility density (miles per square mile x $10^{-1}$) | 0.251 | 2.54 |
|     Interacted with number of bicycles in the household | 0.864 | 2.34 |
| Availability of transit service to work zone | 0.570 | 2.71 |
| Transit access time to stop (minutes x $10^{-1}$) | -0.425 | -5.25 |
| **Zonal demographics and housing cost (including demographic interactions)** | | |
| Absolute difference between zonal median income and household income ($ x $10^{-5}$) | -2.077 | -11.59 |
| Absolute difference between zonal average household size and household size | -0.349 | -5.05 |
| Average of median housing value ($ x $10^{-5}$) | -0.182 | -7.01 |
| **Zonal ethnic composition measure** | | |
| Fraction of Caucasian population interacted with Caucasian dummy variable | 2.836 | 13.82 |
| Fraction of African-American population interacted with African-American dummy variable | 2.736 | 5.18 |
| Fraction of Hispanic population interacted with Hispanic dummy variable | 2.199 | 4.47 |

**TABLE 3. Estimation Results of the Mode Choice Model**

| Variables | Parameter | t-stat |
|---|---|---|
| **Alternative specific constants** | | |
| Auto – Drive alone | 0 | Fixed |
| Auto – Drive with passenger | -3.418 | -16.88 |
| Auto – Passenger | -1.397 | -3.00 |
| Walk | -1.020 | -1.64 |
| Bike | -3.021 | -5.20 |
| Transit | -3.825 | -4.23 |
| **Socio-demographics** | | |
| Number of vehicles per number of licenses – Drive modes | 1.918 | 4.32 |
| Number of bicycles – Bike mode | 0.419 | 7.70 |
| Household size – Passenger and drive passenger modes | 0.170 | 3.04 |
| **Individual level LOS variables (including demographic interactions)** | | |
| Travel time (in minutes) | -0.011 | -1.57 |
| interacted with inflexible work schedule | -0.008 | -1.55 |
| Travel cost (in dollars) | -0.144 | -1.82 |
| **Household level commute-related variables** | | |
| Total drive commute time of all workers (minutes x $10^{-1}$) – Auto modes | 1.336 | 1.60 |
| Standard deviation of the error term common to residential location and mode choice models – Auto modes (negative correlation) | 0.859 | 1.53 |
| **Zonal size and density measures (including demographic interactions)** | | |
| Population density (#households per acre x $10^{-1}$) – Non auto modes | 0.019 | 2.25 |
| Employment density (#employment per acre x $10^{-1}$) – Non auto modes | 0.004 | 2.16 |
| interacted with household income in lowest quartile – Non auto modes | 0.268 | 1.39 |
| **Zonal land-use structure variables** | | |
| Land-use mix – Transit mode | 2.418 | 1.60 |
| **Local transportation network measures (including demographic interactions)** | | |
| Street block density (#blocks/square mile x $10^{-1}$) – Non motorized modes | 0.367 | 2.64 |
| Total length of bikeways within one mile radius (meters x $10^{-5}$) – Bike mode | 1.267 | 1.22 |