Accommodating Immigration Status and Self Selection Effects in a Joint Model of Household Auto Ownership and Residential Location Choice

Chandra R. Bhat *(corresponding author)*
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172
Phone: 512-471-4535, Fax: 512-475-8744
Email: bhat@mail.utexas.edu
and
King Abdulaziz University, Jeddah 21589, Saudi Arabia

Rajesh Paleti
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172
Phone: 512-471-4535, Fax: 512-475-8744
Email: rajeshp@mail.utexas.edu

Ram M. Pendyala
Arizona State University
School of Sustainable Engineering and the Built Environment
Room ECG252, Tempe, AZ 85287-5306
Phone: 480-727-9164; Fax: 480-965-0557
Email: ram.pendyala@asu.edu

Karen Lorenzini
Texas Transportation Institute - Austin
1106 Clayton Lane, Suite 300E, Austin, Texas 78723
Phone: 512-467-0952; Fax: 512-467-8971
Email: k-lorenzini@ttimail.tamu.edu

Karthik C. Konduri
Department of Civil and Environmental Engineering
University of Connecticut
261 Glenbrook Road, Unit 3037
Storrs, CT 06269-3037, USA.
Phone: 860-486-2733; Fax: 860-486-2298
Email: kkonduri@engr.uconn.edu

ABSTRACT
As the proportion of immigrants in the US population continues to rise, it is becoming increasingly important to understand their residential location choices and travel behavior in the travel modeling and transportation policy making arena. This paper presents a joint model of residential location choice and auto ownership that explicitly accounts for immigration status and length of stay in the United States as explanatory variables. Model estimation results using a San Francisco Bay Area subsample of the 2009 National Household Travel Survey (NHTS) show that immigration and length of stay are significant explanatory variables in both residential location choice and auto ownership, with immigrants displaying assimilation effects, *i.e.*, they increasingly resemble non-immigrant households as the length of stay increases. Even after controlling for immigration effects and including residential location choice as an explanatory variable in the auto ownership model, it is found that there are significant self-selection effects that are likely to dampen estimates of the impacts of land use changes on travel behavior in policy forecasts. The paper demonstrates the need to account for immigration variables and self-selection effects in transportation forecasting models that inform policy decisions.

INTRODUCTION

In light of the transformative role that immigration is playing in shaping the US population, there has been considerable interest in studying immigrant activity-travel patterns and ensuring that their mobility needs are met in the policymaking arena (*1,2)*. This paper aims to contribute further to this body of knowledge by *accounting for immigration status* in jointly examining two key facets of behavior that have important implications for activity-travel demand analysis. The two aspects of behavior examined in this paper are vehicle ownership and residential location choice. Both these choices have been the subject of much study in the profession due to their important role in shaping activity-travel patterns of people (*3,4*), and it is therefore of much interest to analyze how immigrants differ from the rest of the population in terms of these choice behaviors. In this paper, both of these choice phenomena are modeled jointly with an aim of accounting for residential self-selection effects that may be present when households choose residential location.

From a substantive standpoint, this study contributes to the understanding of the influence of immigration status on vehicle ownership and residential location choice in four distinct, but inter-related, ways. First, the study expressly controls for the influence of immigration status on vehicle ownership. In addition to providing useful insights into vehicle ownership behavior differences between immigrant households (all members born outside the US), combination households (one or more members born in the US and one or more members born outside the US), and domestic households (all members born in the US), incorporating the immigrant effect on vehicle ownership is important in assessing the impact of the built environment on vehicle ownership. For instance, immigrant households tend to reside in relatively high density central cities or “ethno-burbs”, and also tend to own fewer vehicles. Thus, ignoring the immigrant effect on vehicle ownership can lead to an exaggerated negative effect of high density residential locations on vehicle ownership choices, incorrectly suggesting that neo-urbanist policies aimed at promoting high density developments would have substantial negative effects on household vehicle ownership levels when the “true” negative effect (after accommodating for the effect of immigration status on vehicle ownership) may perhaps be more tempered.

Second, the study controls not only for immigration status effects on vehicle ownership, but also for the length of stay in the U.S. for immigrant and combination households (defined as the length of stay in the US of the person in the household who has been in the US the longest). This allows examining for the presence and magnitude of “assimilation” effects in long-term and short-term travel choices.

Third, in addition to the generic negative immigrant effect on vehicle ownership, there may be variations in the sensitivity of vehicle ownership to built environment measures based on demographic variables and immigrant status, which are tested in this study by including interactions of demographic and immigrant variables with categorical indicator variables of different residential location alternatives (characterized based on the built-up density of the neighborhood).

Fourth, the study considers unobserved self-selection effects by examining if households choose to locate in neighborhoods based on their unobserved (to the analyst) desire or lack of desire to own vehicles. This is achieved through the joint modeling of household residential location choice and household auto ownership choice. In the joint model system estimated for this paper, the presence (or absence) of residential self-selection effects is examined while explicitly accounting for immigration status of the household.

From a methodological standpoint, the study contributes to the literature by presenting a bivariate multinomial probit (MNP) model system, with one MNP model to analyze residential location type choice and another MNP model to examine vehicle ownership choice. In this system, residential location type choice is defined in seven categories based on the built-up density (proxied by the number of households per square mile) of the Census Block in which the household is located. Given that residential density is highly correlated with most measures of urban sprawl, it may be viewed as a reasonable proxy for several built environment characteristics of a neighborhood (*5*). The two MNP models of residential neighborhood type and vehicle ownership are tied together in a bivariate system, and estimated using a Maximum Approximated Composite Marginal Likelihood (MACML) estimation approach (*6*). The system allows for the presence of common unobserved factors (*e.g.*, attitudes and perceptions, individual lifestyle preferences) that affect both vehicle ownership and residential location choice through the presence of error correlations across the choice dimensions of interest. The model system is estimated on a subsample of the 2009 National Household Travel Survey (NHTS) drawn from the San Francisco Bay Area in California where there is a large immigrant population.

The remainder of this paper is organized as follows. The next section offers a discussion on the role of immigration status in shaping vehicle ownership and residential location choice behaviors. The third section describes the data while the fourth section presents the modeling methodology. The fifth section presents model estimation results and the sixth and final section offers concluding thoughts.

IMMIGRATION AND TRAVELER BEHAVIOR

As mentioned in the previous section, there has been considerable interest in analyzing and understanding immigrant travel behavior with an emphasis on identifying their unique mobility needs and informing decision makers in the policymaking arena. From a conceptual standpoint, there are several factors that contribute to the unique residential location and traveler behavior exhibited by immigrants (*7,8*). These include (a) resource-related considerations (such as lower incomes and poor access to mainstream sources of financing because of the absence of a credit history), (b) mobility-related considerations (for example, locating in neighborhoods with more alternative modes of transportation and mixed land-use so that they can have more control over participation in activities through the use of relatively inexpensive forms of transportation such as walking), (c) past travel behavior and cultural norms (for instance, women from certain cultures may be less inclined to drive, or individuals may have not driven in the past, or individuals from some cultures may be more environmentally-friendly, all based on their past behaviors and living circumstances in their originating countries), and (d) self-preservation and identity considerations (for example, locating close to other people of their own culture to feel a sense of closeness/belonging, as well as retaining a sense of security and place in a foreign land).

Studies of residential location choice of immigrant households have shown that racial groups differ with respect to their home ownership rates (*9*) and location choice preferences (*10,11*). In particular, Wilson and Singer (*10*) analyzed 2010 American Community Survey data on foreign born population and found that although 51 percent of immigrants nationwide live in suburban neighborhoods of large metropolitan areas, they are less likely to live in suburban neighborhoods when compared with the overall population in 78 of the 100 largest metro areas. Along similar lines, Logan *et al.* (*11*) studied immigrant enclaves in New York and Los Angeles; they note that, although such central city neighborhoods may be viewed as a temporary base for low income immigrants until they become assimilated into the population, it is also possible that immigrants have an innate preference for residing in these neighborhoods. Contrino and McGuckin (*12*) report that immigrant travel behavior is considerably different from that of native born Americans; although the difference fades as immigrants' period of stay in the US increases, they find that Hispanics are still more likely to use transit than other immigrant groups even after 20 years in the US. Ma and Srinivasan (*13*) estimate an ordered response probit model of auto ownership that explicitly includes explanatory variables representing the immigration status and characteristics of the household. Their results show a significant impact of duration of stay and time period of entry on the propensity for car ownership, with a decreasing influence of immigrant status variables as an individual assimilates over a period of time. This study, however, does not account for residential self-selection effects wherein auto ownership and choice of residential location type constitute a joint choice phenomena representative of a household's lifestyle preference.

A number of studies have examined mode usage patterns of immigrant populations to better understand the underlying factors motivating mode choices made by such populations. Rosenbloom (*14*), in assessing the transit markets of the future, notes that immigrants constitute one of the key markets for transit services and that transit services need to be customized to meet their unique mobility needs. However, one challenge that arises is that as immigrants increasingly assimilate into the general population, they may decrease their use of transit modes. Blumenberg (*1)* notes that the rising tide of immigrants in the nation increases the use of all travel modes. Immigrants are more likely to be transit users, at least for a while until they assimilate and become more similar to the general population in their travel choices. The reasons for higher transit use among immigrants include (a) immigrants initially reside in high density neighborhoods which are better served by transit, (b) many immigrants are from nations where auto use is less than that in the United States, and (c) legal restrictions and processes which must be completed prior to obtain a valid driver's license (*8*). However, Blumenberg (*1*) also finds that, despite assimilation effects, immigrants show a higher rate of usage of alternative modes of travel such as non-motorized travel, car pooling, and transit even after a long period of residence in the US. Similar findings are also reported by Chatman (*2*) who find that immigrants carpool and use transit more than others even after living in the United States for 20 years. It is interesting to note that not only immigrants, but native-born US citizens exhibit a change in travel behavior after experiencing a temporary residence abroad (*15*) suggesting that people tend to adopt the travel behavior patterns of the location in which they are placed.

Tal and Handy (*16)* also report considerable differences in immigrant travel behavior in an analysis of the 2001 National Household Travel Survey (NHTS) data set. They note that while their analysis clearly points to differences in behavior, these differences do not necessarily imply a causal effect. They argue that it is not clear why immigrant status affects travel behavior; perhaps immigrant status variables represent a number of factors such as needs, preferences, cultural differences, and prior experiences that are the true determinants of travel choices. Smart (*7*) notes that immigrants are more likely to use the bicycle mode for travel even after controlling for other socio-demographic characteristics. There is a large and significant "immigrant effect" that varies somewhat by country of origin with those from East and Southeast Asia more likely to use the bicycle in comparison to other immigrant groups. Valenzuela *et al.* (*17*) note that immigrant groups are more likely to rely on informal travel arrangements; this may motivate immigrants to reside in neighborhoods with a high concentration of immigrants of similar heritage and may also decrease the need for owning a personal vehicle (as immigrants share rides and use other informal travel means). Liu and Painter (*18*) report similar findings in an analysis of travel behavior of Latino immigrants that examines the role of ethnic concentration and ethnic employment on traveler choices.

It is clear from the literature that immigrant status is strongly associated with travel behavior and mode usage. The evidence in the literature suggests that immigrants self-select when choosing residential location; they tend to locate themselves in ethnic enclaves, in higher density central city locations, and generally more transit- and pedestrian- oriented neighborhoods. There is a vast body of literature that has found higher density land use associated with lower levels of car ownership and use, and higher levels of walking, bicycling, and transit use (*5,19,20*). While these and many other studies have shown the association between land use density and travel behavior, they also note that there may be self-selection effects wherein people with a proclivity for a transit- and pedestrian-oriented lifestyle choose to locate themselves in neighborhoods conducive to such a lifestyle. There have been a number of studies examining residential self-selection effects (*21,22*) with a goal of more accurately assessing the impacts of land use on travel behavior. If people of a certain type (characterized by their attitudes, perceptions, and preferences) self-locate in high-density neighborhoods, then the actual impact of land use policy measures on travel behavior is likely to be less than predicted if such sorting effects are ignored.

The issue of self-selection is further complicated in the context of immigrant status. Two phenomena (of interest in the context of this paper) may be at play when considering the land use-travel behavior relationship. The true impacts of land use on vehicle ownership and use may be dampened due to two inter-related factors. First, immigrants tend to use transit and alternative modes more than others regardless of where they live. In other words, the apparent impact of land use on travel behavior may, at least in part, be due to immigrant status (rather than land use per se). Second, immigrants have an innate preference for lifestyles that are characterized by alternative mode use and lower car ownership. As such, there may be unobserved self-selection effects arising from immigrant residential location choice behavior consistent with that postulated in previous studies of residential self-selection. Thus, there may be both observed and unobserved immigrant status effects that should be reflected in models of residential location choice and auto ownership (both of which are key variables in transportation forecasting models). A study that attempts to examine a multitude of travel choices is that by Beckman and Goulias (*23*), who model travel time, departure time, and mode choice using latent class cluster analysis. They used residential location, car ownership, immigration status, and several other key demographic variables as determinants of the latent population class or segment to which a household belonged. Subsequently, commute behavioral characteristics were modeled conditional on the latent class to which the household was assigned. . Their study does not, however, account for self-selection effects. This paper contributes further to this body of work by examining residential location and car ownership of immigrant populations in a joint modeling framework that explicitly accounts for residential self-selection effects while controlling for immigrant status.

DATA

The data used in this study is derived from the 2009 National Household Travel Survey (NHTS), which provides detailed socio-economic, demographic, vehicle ownership and use, and travel information for a large sample of households in the United States. In addition, the data set includes information about the type and nature of the location in which the household respondent resides. For the current study, the subsample residing in the San Francisco Bay Area is extracted and used for analysis and modeling purposes. The subsample from this geographic region includes many immigrants, both well-established as well as new arrivals, that is well-suited for examining the influence of immigrant status on residential location choice and auto ownership. After extensive data cleaning, the San Francisco Bay Area data set is comprised of 3,335 households that provided complete information on a host of socio-economic, demographic, and travel variables of importance to this study.

An examination of the socio-economic and demographic characteristics of the subsample of households showed that the sample is reasonably representative of the region and suitable for analyzing residential location and auto ownership choices of households. For the sake of brevity, a complete socio-economic and demographic profile of the sample is not furnished here. With respect to vehicle ownership (a variable of interest in this study), it was found that 4.8 percent of the households reported zero vehicle ownership, 26 percent reported owning one vehicle, 43 percent reported owning two vehicles, 17 percent reported owning three vehicles, 6 percent reported owning four vehicles, and fewer than 3 percent reported owning five or more vehicles.

Immigration status is reported at the level of the individual person based on whether the person was born in the United States or not. Based on the individual immigrant status, households were classified into three immigrant categories. About 9 percent of households included only immigrant household members and may be considered as pure immigrant households. About 17.5 percent of households included a combination of immigrant and non-immigrant household members, and may be considered as combination households. Finally, about 74 percent of households were purely non-immigrant households. The combination households are largely those comprising immigrant parents with native-born children. A comparison of these proportions against those in the 2009 American Community Survey (ACS) data suggests that there is a modest under-representation of immigrant households in the NHTS data set. However, this does not present a problem within the context of the modeling effort of this study, and an exploration of the implications of under-representation of immigrants in survey samples on travel demand forecasts remains an exercise for future research.

The survey data set included information about the density of the census block in which the household resides. As the analysis of land use-travel behavior relationships continues to rely heavily on measures of density, residential location choice is categorized by population density within this study. Specifically, households are considered as choosing among seven possible residential location alternatives (census block types). The distribution of households across these seven alternatives is as follows:

- 0-99 households per square mile: 4.0%
- 100-499 households per square mile: 9.7%
- 500-999 households per square mile: 8.9%
- 1000-1999 households per square mile: 19.7%
- 2000-3999 households per square mile: 35.7%
- 4000-9999 households per square mile: 16.3%
- ≥ 10000 households per square mile: 5.8%

MODELING METHODOLOGY

This section presents the modeling methodology adopted in this paper.

Model Framework

The methodology in this paper is concerned with the estimation of joint models of any number of nominal variables. Let g be the index for the nominal variables ($g = 1, 2, \ldots, G$). In the current empirical context, G=2, where the two nominal variables are residential location (g=1) and household auto ownership (g=2) respectively. Also, let I_g be the number of alternatives corresponding to the g^{th} nominal variable ($I_g \geq 2$) and let i_g be the corresponding index (i_g = 1, 2, 3, …, I_g). In the current empirical context, $I_1 = 7$ (seven residential location alternatives defined based on the built-up density of the census block) and I_2 = 5 (household auto ownership ranging from 0 to 4 or more).

Under the usual assumptions of random utility theory, the utility associated with alternative i_g of the g^{th} nominal variable may be written as:

$$U_{gi_g} = \boldsymbol{\beta}'_g \boldsymbol{x}_{gi_g} + \varepsilon_{gi_g}, \tag{1}$$

where $\boldsymbol{x}_{gi_g}$ is a (K_g×1)-column vector of determinant attributes. Note that, in addition to demographic and immigrant status variables, the residential location dummy variables that form the observed dependent variables for the residential location choice model (g=1) appear in the explanatory attribute vector in the utilities of the car ownership model (g=2). $\boldsymbol{\beta}_g$ is a column vector of corresponding coefficients, and ε_{gi_g} is a normal error term. Let the variance-covariance matrix of the vertically stacked vector of errors $\varepsilon_g [= (\varepsilon_{g1}, \varepsilon_{g2}, ...\varepsilon_{gI_g})']$ be $\boldsymbol{\Omega}_g$. Assuming that the household chooses the alternative m_g, all the utility differences with respect to this chosen alternative m_g must be less than zero. These conditions can be denoted numerically as:

$$U_{gi_g} - U_{gm_g} < 0 \,\forall\, i_g \neq m_g \tag{2}$$

Let $y^*_{gi_g m_g} = U_{gi_g} - U_{gm_g} \, (i_g \neq m_g)$, and $\boldsymbol{y}^*_{\boldsymbol{g}}$ be the stacked vector of the latent utility differentials $\boldsymbol{y}^*_{\boldsymbol{g}} = \left[\left(y^*_{g1m_g}, y^*_{g2m_g}, ..., y^*_{gI_g m_g} \right)'; i_g \neq m_g \right]$. $\boldsymbol{y}^*_{\boldsymbol{g}}$ has a mean vector of $\boldsymbol{B}_{\boldsymbol{g}} \, (\boldsymbol{\beta}'_1 \boldsymbol{z}_{\boldsymbol{g1m_g}}, \boldsymbol{\beta}'_1 \boldsymbol{z}_{\boldsymbol{g2m_g}}, ...\boldsymbol{\beta}'_1 \boldsymbol{z}_{\boldsymbol{gI_g m_g}})'$, where $\boldsymbol{z}_{\boldsymbol{gi_g m_g}} = \boldsymbol{x}_{\boldsymbol{gi_g}} - \boldsymbol{x}_{\boldsymbol{gm_g}}, i_g = 1, 2, ..., I_g;\; i_g \neq m_g$ and a covariance matrix given by $\boldsymbol{\Sigma}^*_{\mathbf{g}} = \boldsymbol{M}_{\boldsymbol{g}} \boldsymbol{\Omega}_{\mathbf{g}} \boldsymbol{M}'_{\boldsymbol{g}}$, where $\boldsymbol{M}_{\mathbf{g}}$ is an $(I_g - 1) \times I_g$ matrix that corresponds to an $(I_g - 1)$ identity matrix with an extra column of –1's added as the m_g^{th} column. Thus, we can write

$$\boldsymbol{y}^*_{\boldsymbol{g}} \sim N\left(\boldsymbol{B}_{\boldsymbol{g}}, \boldsymbol{\Sigma}^*_{\mathbf{g}}\right), \tag{3}$$

Now, for the 2 nominal variables, consider the stacked $\widetilde{G} = (I_1 + I_2 - 2)$ vector $\boldsymbol{y}^* = \left[\left(\boldsymbol{y}^{*'}_1, \boldsymbol{y}^{*'}_2 \right)' \right]$, each of whose element vectors is formed by differencing utilities of alternatives from the chosen alternative m_g for the g^{th} nominal variable. Then, we may write: $\boldsymbol{y}^* \sim N(\boldsymbol{B}, \boldsymbol{\Sigma}^*)$, where $\boldsymbol{B} = \left(\boldsymbol{B}'_1, \boldsymbol{B}'_2 \right)'$ and $\boldsymbol{\Sigma}^*$ is a $\widetilde{G} * \widetilde{G}$ matrix as follows:

$$\boldsymbol{\Sigma}^* = \begin{bmatrix} \boldsymbol{\Sigma}^*_1 & \boldsymbol{\Sigma}^*_{12} \\ \boldsymbol{\Sigma}^*_{21} & \boldsymbol{\Sigma}^*_2 \end{bmatrix} \tag{4}$$

The off-diagonal elements in $\boldsymbol{\Sigma}^*$ capture the dependencies across the utility differentials of the two nominal variables, the differential being taken with respect to the chosen alternative for each nominal variable. Let $\boldsymbol{\theta}$ be the collection of parameters to be estimated:

$\boldsymbol{\theta} = [\boldsymbol{\beta_1}, \boldsymbol{\beta_2}; \text{Vech}(\boldsymbol{\Sigma}^*)]$, where Vech($\boldsymbol{\Sigma}$) represents the vector of upper triangle elements of $\boldsymbol{\Sigma}$. Then the likelihood function for a household may be written as:

$$L(\boldsymbol{\theta}) = F_{\tilde{G}}[-\tilde{\boldsymbol{B}}, \boldsymbol{\Sigma}^*], \tag{5}$$

where $F_{\tilde{G}}(.,.)$ is the $\tilde{G} = (I_1 + I_2 - 2)$-dimensional normal cumulative distribution function.

The above likelihood function involves the evaluation of a $\tilde{G} = (I_1 + I_2 - 2)$-dimensional integral for each individual household, which can be very computationally expensive if each nominal variable can take a large number of values. Therefore, in this study, the Maximum Approximate Composite Marginal Likelihood (MACML) approach (*6*) is used. In this approach, the likelihood function only involves the computation of univariate and bivariate cumulative distributive functions.

The MACML Estimation Approach

The MACML estimation approach proceeds by working with pairs of nominal variables instead of all nominal variables together. As there are only two nominal variables in the current empirical context, the pairwise likelihood approach of the MACML method and the full likelihood approach are essentially the same. Specifically:

$$L_{ML}(\boldsymbol{\theta}) = L_{CML}(\boldsymbol{\theta}) = \Pr(d_{i_1} = m_1, d_{i_2} = m_2) \tag{6}$$

where d_{i_g} is an index for the household's choice for the g^{th} nominal variable, and m_g is the actual chosen alternative for the g^{th} nominal variable. However, in cases where there are several alternatives for one or more nominal variables, the dimension $\tilde{G} = (I_1 + I_2 - 2)$ can be high. In that case, an analytic approximation to the multivariate integral is used to evaluate the high dimensional integral (*6*). The resulting maximum approximate composite marginal likelihood (MACML) estimation is solely based on bivariate and univariate cumulative normal computations.

A couple of issues arise in the specification and estimation of joint multivariate discrete choice models with error correlations. The first is that appropriate normalizations and identification restrictions need to be placed on parameters to ensure that coefficients and error covariance parameters can be uniquely identified. In addition, it is necessary to ensure that the covariance matrices of utility differences that are constructed during the estimation process are all positive definite. A detailed discussion of these two issues is beyond the scope of this paper. For details on the normalization procedures to ensure identification and positive definiteness, the reader is referred to Bhat (*6*).

ESTIMATION RESULTS

This section presents a detailed discussion of the model estimation results.

Residential Location Choice Component

Estimation results for the residential location choice component of the joint model system are presented in Table 1. Household density of fewer than 99 households per square mile (at the census block level) is the base alternative. An examination of the alternative specific constants suggests that households generally prefer to locate in medium-density neighborhoods as opposed to very low density or very high density areas. Asian households, in comparison to all other races, are likely to locate in very high density areas, clearly suggesting that this particular ethnic group has a unique residential location preference. The presence of children is associated with a lower likelihood of residing in high density areas, presumably because households with children seek to reside in suburban neighborhoods which are often thought to have good schools and more open space. Households that have an individual with a prolonged medical condition are more likely to reside in high density areas, presumably because they want easier access to destinations and opportunities (including medical care). Results show that households with larger number of individuals who are self-employed or have more than one job prefer to locate in low density areas; the reasons for this are not readily apparent and merit further investigation. Home ownership is associated with living in lower density neighborhoods, which is consistent with expectations as households tend to purchase homes in the affordable suburban areas that are often viewed as being more family-oriented than dense developments or apartment complexes.

Income has a significant influence on household residential location choice. Lower income households appear to be more likely to reside in higher density areas relative to higher income households. The latter group may prefer to reside in larger households in the suburban lower density locations and are able to afford the transportation costs of residing away from the city. Even after controlling for all of these demographic variables, immigration variables are significant in predicting household residential location. Immigrant households and combination households are more likely to reside in dense census blocks, clearly suggesting that immigrants seek to live in higher density neighborhoods – a finding that is consistent with the large body of literature on the subject – except that this analysis has shown that this relationship exists even after controlling for other socio-demographic characteristics (*10*). However, it is also seen that, as the length of stay in the United States increases, the proclivity to reside in higher density areas decreases – suggesting that immigrant households begin to mimic the residential location choice patterns of non-immigrant households (*i.e.*, reside in suburbs) as they assimilate.

Auto Ownership Choice Component

Estimation results for the auto ownership model component are shown in Table 2. The zero-vehicle choice is the base alternative in this model. It can be seen from an examination of the alternative specific constants that households generally tend to own one or two cars (as opposed to zero cars or three or more cars) consistent with expectations. Higher levels of car ownership are associated with larger numbers of adults in the household. Households with individuals who are likely to have mobility limitations (senior adults, adults with medical condition) are more likely to own zero cars as evidenced by the negative coefficients on these variables for all non-zero auto ownership levels. Households with low education attainment (high school or less) show a proclivity to own cars, presumably because workers in these households need the car to get to work. As the educational attainment increases to completion of college degrees, households have a lower likelihood of owning three or more vehicles. This may be indicative of such households locating in neighborhoods with good transit access or being (environmentally) conscious of their travel choices and choosing to own fewer vehicles. In terms of work characteristics, households with a larger number of self-employed individuals or individuals with more than one job are likely to exhibit higher car ownership levels; however, this may simply be due to the fact that such households tend to have larger household sizes and hence exhibit higher levels of car ownership. As the mean distance to work (across workers in the household) increases, car ownership levels increase; this is consistent with expectations as people may prefer to drive and have access to their own personal vehicle to navigate longer commutes.

The impact of household income on vehicle ownership is perfectly consistent with expectations. As incomes rise, so do car ownership levels with very clear trends in the magnitudes of coefficients along expected lines. Home ownership is associated with higher levels of car ownership, which is again consistent with expectations as home owners are likely to be higher income households in suburban locations with children – all of which contribute to higher levels of auto ownership.

Immigrant variables have a significant impact on auto ownership choice even after controlling for other demographic variables. Immigrant households are less likely to exhibit higher levels of car ownership as evidenced by negative coefficients in the three-vehicle and four-or-more vehicle categories. However, as the length of stay in the US increases, there is a positive inclination to acquire vehicles (across all vehicle ownership levels) relative to the zero-vehicle ownership category. This is suggestive of an assimilation effect over time, although the assimilation period differs substantially across vehicle ownership levels. The assimilation process seems to be reasonably fast at lower car ownership levels with the preference for one vehicle auto ownership of immigrant households catching up with corresponding preference of non-immigrant households within six years of moving to the US. For higher auto ownership levels, the assimilation process is quite slow taking more than 40 years. There is a significant interaction effect between immigrant status, length of stay, and income level. Although high incomes are generally associated with higher auto ownership levels, immigrants at this income level who have stayed less than 10 years in the country show a disinclination to own vehicles (at all vehicle ownership levels). Thus, in the case of immigrants, the length of stay appears to be an important driver of vehicle ownership; in the initial period (less than 10 years), even high income immigrants appear to mimic the travel choices that they may have been accustomed to in their original country.

The direct effect of residential location on auto ownership is quite substantial. Households residing in the lowest density neighborhoods exhibit a higher inclination towards owning four or more vehicles, presumably because alternative modes of travel may not be available in such neighborhoods and/or because households that locate in such low density areas prefer auto-oriented lifestyles. Households in the highest density categories, on the other hand exhibit a lower inclination to own cars relative to households in moderately dense neighborhoods. These neighborhoods are likely to have good transit service and be pedestrian and bike friendly; moreover, households that locate in such dense neighborhoods are likely to favor a lifestyle with limited automobile dependency and use. Results show that residential location choice significantly impacts auto ownership choice even after controlling for a number of other demographic variables.

Self-Selection Effects

The significance of self-selection effects in the joint residential location-vehicle ownership choice model system may be determined by examining the covariance matrix and comparative goodness-of-fit statistics depicted in Table 3. The table presents estimation results corresponding to the covariance matrix of utility differences of residential location and auto ownership alternatives. The elements in the block diagonal sub-matrices (shaded in gray) correspond to each of the MNP model components of residential location and auto-ownership. The positive estimate of 0.5952, which is significantly different from 1, shows that the variance of unobserved factors that influence the utility associated with choosing moderately low-dense (500-999 households per square mile) neighborhoods is lower than that of other residential location alternatives. An independent MNP model that does not accommodate heterogeneity in the unobserved factors would have imposed identical variance distribution assumptions across all utilities. In the auto ownership component block sub-matrix, the covariance estimate of 0.9832 suggests that the unobserved factors which influence the utilities of owning one vehicle or owning two vehicles are positively correlated with one another. An independent MNP model system would have ignored such correlations across unobserved factors. It is possible that attitudes, lifecycle stage, and mobility needs (unobserved factors) that motivate a household to acquire one vehicle also motivate a household to acquire two vehicles (recall that most households in the sample have one or two vehicles). However, as car ownership moves to the three and four or more levels, then other motivating factors may be at play.

The non-zero non-block elements in the covariance matrix (unshaded region) capture the self-selection effects (that is, the jointness arising from error correlations across the two choice dimensions). The t-statistics for these elements signify whether the parameter is significantly different from zero. Even after controlling for the direct effect of residential density and immigration status (along with several other demographics) on auto ownership, estimation results show that there is a statistically significant error correlation indicative of unobserved self-selection effects. Specifically, the estimates of 0.0747 and 0.0883 suggest that unobserved factors contributing to a choice of low density residential area also contribute positively to the highest level of auto ownership. These households are likely to be auto-oriented households that prefer a lifestyle involving automobile ownership and use. In other words, unobserved lifestyle preferences positively contribute both to high auto ownership and to the choice of a low density residential area.

The negative covariance parameters in the higher residential density alternatives are also indicative of self-selection effects (-0.0972, -0.1204, -0.2388, -0.1702). These parameters suggest that unobserved factors that motivate households to choose higher density neighborhoods for their residential location are negatively correlated with unobserved factors that contribute towards household vehicle ownership. This finding is consistent with expectations and the literature on residential self-selection. Households that have a preference towards alternative mode use (such as transit and walk/bike) may seek to live in high density neighborhoods and thus own fewer vehicles. Households that have a preference for an auto-oriented lifestyle may seek to acquire more cars and live in low density neighborhoods. In other words, the unobserved factors contributing to high density residential location are negatively correlated with unobserved lifestyle preferences contributing to auto ownership. In a model system that ignores jointness in residential location choice and auto ownership, such self-selection effects would be ignored potentially leading to erroneously amplified predictions of the impact of land use densification on auto ownership.

The model estimation results demonstrate that immigration status is an important predictor of residential location choice and auto ownership. The results also suggest that self-selection effects are present and significant. To further establish the importance of considering immigration status and self-selection effects, the data fit of the joint model presented in this paper is compared against that of a series of other models that ignore one or both of these effects. Three additional models which ignored either one or both of the immigration status and self-selection effects were estimated and compared against the joint model presented in this paper. The lower section of Table 3 presents the log-likelihood ratio (LR) test statistics obtained when comparing the joint model against each of the other three models. It can be seen from the last column of the table that the LR statistic is consistently greater than the critical chi-squared value, demonstrating the significant improvements in data fit as one progressively accounts for immigration and self-selection effects.

CONCLUSIONS

This study aims to explore the influence of immigrant status on residential location and auto ownership choices by explicitly including immigration variables as explanatory factors in the choice models. Further, the potential presence of assimilation effects is explored by considering the length of stay in the United States as an explanatory variable. A joint bivariate multinomial probit (MNP) model of residential location choice and auto ownership choice that includes immigration-related explanatory variables is estimated using the maximum approximate composite marginal likelihood (MACML) method on a San Francisco Bay Area subsample of the 2009 National Household Travel Survey (NHTS) data set. The joint bivariate model is capable of capturing self-selection effects that may exist even after accounting for the influence of immigration status *and* including residential location choice as an explicit explanatory variable in the auto ownership model. Thus, this paper offers a methodological contribution by presenting a computationally tractable model formulation and estimation approach capable of capturing self-selection effects in multi-dimensional choice models.

Model estimation results show that immigration status is an important explanatory variable for both residential location choice and auto ownership choice, even after controlling for a number of other socio-economic and demographic variables. Immigrants appear to favor residing in higher density neighborhoods and owning fewer cars relative to non-immigrants, a finding that is essentially consistent with previous literature. Study results also show that immigrants tend to assimilate over time, with the length of stay in the United States contributing towards making immigrants similar to non-immigrant (*i.e.*, native born) households in terms of residential location and auto ownership choices. However, assimilation periods can be quite long (on the order of 40 years) for certain choice alternatives (*e.g.*, owning four or more cars), suggesting that immigrant households tend to retain differences vis-à-vis non-immigrant households for a long period of time (lasting as long as one or two generations). Model estimation results also show significant self-selection effects with significant correlations in unobserved factors that affect both residential location choice and auto ownership choice.

Findings of this study have important planning and policy implications. The study shows that the inclusion of immigrant-status variables is critical to developing travel demand forecasts that account for the diversity of the US population. Socio-economic and demographic forecasting models that provide inputs to travel demand models should include the ability to forecast immigrant presence in the population so that travel forecasts explicitly account for their behavior. Also, model estimation results show that unobserved lifestyle preferences and household attitudes towards vehicle ownership (and use) motivate households to self-select into neighborhoods whose characteristics are consistent with those preferences. Ignoring such self-selection effects leads to erroneous model parameter estimates that would result in over-estimating the impact of land use changes on travel behavior. This finding has important implications in the policy arena where land use interventions (commonly including densification and diversification of land uses) are often proposed to address energy and environmental concerns. It is likely that the true impacts of land use policies on travel behavior are lower than what one would have estimated had the effects of immigrant status and residential self-selection been ignored. In addition, the presence of lingering differences in location and travel behavior characteristics among immigrants when compared with the general population (even after accounting for assimilation effects) constitutes an opportunity for transportation planners to design and provide alternative residential and transportation choices that meet the lifestyle preferences of immigrants and result in a smaller energy and transportation-related carbon footprint.

REFERENCES

1. Blumenberg, E. Moving in and moving around: immigrants, travel behavior, and implications for transport policy. *Transportation Letters*, Vol. 1, No. 2, 2009, pp. 169-180.
2. Chatman, D.G., and N. Klein. Immigrants and travel demand in the United States: Implications for transportation policy and future research. *Public Works Management & Policy*, Vol. 13, No. 4, 2009, pp. 312-327.
3. Mokhtarian, P.L., and X. Cao. Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B*, Vol. 42, No. 3, 2008, pp. 204-228.
4. Bhat, C.R., and J.Y. Guo. A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B*, Vol. 41, No. 5, 2007, pp. 506-526.
5. Brownstone, D., and T. Golob. The impact of residential density on vehicle usage and energy consumption. *Journal of Urban Economics*, Vol. 65, 2009, pp. 91-98.
6. Bhat, C.R. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B*, Vol. 45, No. 7, 2011, pp. 923-939.
7. Smart, M. US immigrants and bicycling: Two-wheeled in Autopia. *Transport Policy*, Vol. 17, No. 3, 2010, pp. 153-159.
8. Blumenberg, E., and A. Norton. Falling immigration rates mean falling transit ridership. *Access*, Vol. 37, 2010, pp. 10-16.
9. Painter, G., S. Gabriel, and D. Myers (2001) Race, immigrant status, and housing tenure choice. *Journal of Urban Economics*, 49, 150-167.
10. Wilson, J.H., and A. Singer. Immigrants in 2010 Metropolitan America: A Decade of Change. State of Metropolitan America, Brookings Institution, Washington, D.C., 2010.
11. Logan, J.R., W. Zhang, and R.D. Alba. Immigrant enclaves and ethnic communities in New York and Los Angeles. *American Sociological Review*, Vol. 67, No. 2, 2002, pp. 299-322.
12. Contrino, H., and N. McGuckin. Travel demand in the context of growing diversity. *TR News*, 2009, 264.
13. Ma, L., and S. Srinivasan. Impact of individuals' immigrant status on household auto ownership. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2156, 2010, pp. 36-46.
14. Rosenbloom, S. *Transit Markets of the Future. The Challenge of Change*. Transit Cooperative Research Program, Report 28. National Academy Press, Washington, D.C., 1998.
15. Burbidge, S.K. Foreign living experience as a predictor of domestic travel behavior. *Journal of Transport Geography*, Vol. 22, 2012, pp. 199-205.
16. Tal, G., and S. Handy. Travel behavior of immigrants: An analysis of the 2001 National Household Transportation Survey. *Transport Policy*, Vol. 17, No. 2, 2010, pp. 85-93.

17. Valenzuela, A., Jr., L. Schweitzer, and A. Robles. Camionetas: Informal travel among immigrants. *Transportation Research Part A*, Vol. 39, No. 10, 2005, pp. 895-911.
18. Liu, C.Y., and G. Painter. Travel behavior among Latino immigrants: The role of ethnic concentration and ethnic employment. *Journal of Planning Education and Research*, Vol. 32, No. 1, 2012, pp. 62-80.
19. Bhat, C.R., S. Sen, and N. Eluru. The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B*, Vol. 43, No. 1, 2009, pp. 1-18.
20. Rajamani, J., C.R. Bhat, S.L. Handy, G. Knaap, and Y. Song. Assessing the impact of urban form measures on nonwork trip mode choice after controlling for demographic and level-of-service effects. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1831, 2004, pp. 158-165.
21. Pinjari, A.R., R.M. Pendyala, C.R. Bhat, and P.A. Waddell. Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. *Transportation*, Vol. 34, No. 5, 2007, pp. 557-573.
22. Bhat, C.R., and N. Eluru. A copula-based approach to accommodate residential self-selection effects in travel behavior modeling, *Transportation Research Part B*, Vol. 43, No. 7, 2009, pp. 749-765.
23. Beckman, J., and K.G. Goulias. Immigration, residential location, car ownership, and commuting behavior: a multivariate latent class analysis from California. *Transportation*, Vol. 35, No. 5, 2008, pp. 655-671.

TABLE 1 Residential Location Component

Residential Location Component (Density less than 99 housing units per square mile is the base)	**Alternative (defined by range of the density of the Census block)**											
	(100-499)		**(500-999)**		**(1,000-1,999)**		**(2,000-3,999)**		**(4,000-9,999)**		**(10,000 and more)**	
Variables	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Constant	0.3478	9.06	0.9523	10.42	0.8889	13.04	1.1629	16.98	1.1880	15.07	0.7369	9.39
Household Demographics												
Race of household (Base: All other races)												
Asian											0.2081	2.39
Presence of children aged between 11-15 years									-0.2395	-3.62	-0.2395	-3.62
Presence of people with prolonged (> 4 years) medical condition							0.1436	2.77	0.1436	2.77	0.1436	2.77
Work Characteristics												
Number of self employed individuals			-0.1485	-3.43	-0.1485	-3.43	-0.1485	-3.43	-0.1485	-3.43	-0.1485	-3.43
Number of individuals with more than one job			-0.1043	-1.62	-0.1043	-1.62	-0.1043	-1.62	-0.1043	-1.62	-0.1043	-1.62
Household Income (Base is Greater than $75K)												
Less than $20K									0.1519	2.37	0.1519	2.37
Between $20K-$45K									0.1382	2.62	0.1382	2.62
Between $45K-$60K									0.1253	2.01	0.1253	2.01
Between $60K-$75K									0.1804	2.68	0.1804	2.68
Housing Tenure: Own House			-0.2441	-3.65	-0.2441	-3.65	-0.2441	-3.65	-0.8474	-11.64	-0.8474	-11.64
Immigration Variables												
Immigrant households							0.3292	3.49	0.3621	3.69	0.3621	3.69
Combination households							0.3188	3.95	0.3227	3.99	0.3227	3.99
Length of stay in the US (in years)							-0.0088	-4.66	-0.0088	-4.66	-0.0090	-3.47

TABLE 2 Auto Ownership Component

Auto Ownership Component (0 vehicles is the base)	Alternatives (0 vehicles is base)							
	1 vehicle		2 vehicles		3 vehicles		4 or more vehicles	
Variables	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Constant	0.6518	4.96	0.2561	1.06	-1.6138	-7.88	-2.4713	-11.59
Household Demographics								
Number of adults			0.1772	2.15	0.6309	8.85	0.8240	11.05
Presence of senior adults (aged>65 years)	-0.4895	-4.29	-0.5307	-4.65	-0.7915	-6.26	-0.9031	-6.54
Presence of people with prolonged (> 4 years) medical condition	-0.4765	-4.00	-0.5141	-4.37	-0.5492	-3.91	-0.5590	-3.46
Highest Education Status (Base is Some college)								
High school or less	0.2033	1.50	0.2033	1.50	0.2033	1.50	0.2033	1.50
Bachelor's degree					-0.2373	-3.26	-0.3086	-3.48
Post Graduate degree					-0.3099	-4.25	-0.4527	-5.25
Work Characteristics								
Number of self employed individuals			0.0477	1.80	0.1033	1.84	0.1033	1.84
Number of individuals with more than one job							0.1842	2.07
Mean Distance to work (in miles)	0.0289	3.25	0.0309	3.56	0.0357	4.07	0.0363	4.08
Household Income								
Between $20K-$45K	0.5235	4.45	0.5659	4.80	0.5659	4.80	0.5659	4.80
Between $45K-$60K	0.7197	4.08	0.8111	4.47	1.0082	5.13	1.1496	5.03
Between $60K-$75K	0.8216	3.59	0.9419	4.14	1.2603	5.16	1.3451	5.03
Greater than $75K			0.1725	2.08	0.5913	5.76	0.8136	5.92
Housing Tenure: Own House	0.7686	7.04	0.8563	7.29	1.0931	7.95	1.0874	7.09
Immigration Variables								
Immigrant households			-0.0710	-1.80	-0.5382	-4.26	-0.7354	-3.90
Length of stay in the U.S.	0.0106	2.99	0.0122	3.47	0.0132	3.48	0.0162	4.08
High income (>-$75K) Immigrants with length of stay less than 10 years in the U.S.	-0.8447	-4.80	-0.8447	-4.80	-0.8447	-4.80	-0.8447	-4.80
Residential Location (Density in housing units per square mile)								
0-99							0.5396	4.01
4,000-9,999							-0.5701	-3.23
10,000 and more	-0.8775	-6.08	-1.0386	-6.92	-1.3663	-5.83	-1.3663	-5.83

TABLE 3 Covariance Matrix of Joint Residential Location and Auto Ownership Model & Goodness of Fit Statistics

		Res (100-499)	Res (500-999)	Res (1,000-1,999)	Res (2,000-3,999)	Res (4,000-9,999)	Res (10,000 & more)	Auto 1	Auto 2	Auto 3	Auto 4 or more
Res	**(100-499)**	1									
Res	**(500-999)**	0.5	**0.5952 (8.38)**								
Res	**(1,000-1,999)**	0.5	0.5	1							
Res	**(2,000-3,999)**	0.5	0.5	0.5	1						
Res	**(4,000-9,999)**	0.5	0.5	0.5	0.5	1					
Res	**(10,000 and more)**	0.5	0.5	0.5	0.5	0.5	1				
Auto	**1**	0	0	0	0	**-0.0972 (-1.61)**	0	1			
Auto	**2**	0	0	0	0	**-0.1204 (-1.95)**	0	**0.9832 (30.2)**	1		
Auto	**3**	0	0	0	0	**-0.2388 (-3.7)**	**-0.1702 (-2.00)**	0.5	0.5	1	
Auto	**4**	**0.0747 (1.66)**	**0.0883 (2.87)**	0	0	0	0	0.5	0.5	0.5	1

Goodness of Fit Statistics

Immigration Status Considered?	Self-Selection Considered?	Log-likelihood	Number of variables	Degrees of Freedom	Critical Chi Square Statistic	LR Test Statistic
Yes	Yes	-9016.27	89	--	--	--
Yes	No	-9065.23	81	8	15.5	29.6
No	Yes	-9080.03	74	15	25.0	127.5
No	No	-9115.65	66	23	35.2	198.8