

The Modeling of Household Vehicle Type Choice Accommodating Spatial Dependence Effects

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
E-mail: rajeshp@mail.utexas.edu

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
E-mail: bhat@mail.utexas.edu

and

King Abdulaziz University, Jeddah 21589, Saudi Arabia

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

Konstadinos G. Goulias

University of California
Department of Geography
Santa Barbara, CA 93106-4060
Phone: 805-308-2837, Fax: 805-893-2578
Email: goulias@geog.ucsb.edu

ABSTRACT

Household vehicle ownership and fleet composition are choice dimensions that have important implications for policy making, particularly in the energy and environmental sustainability arena. In the context of household vehicle ownership and type choice, it is conceivable that there are substantial spatial interaction effects due to both observed and unobserved factors. This paper presents a multinomial probit model formulation that incorporates spatial spillover effects arising from both observed and unobserved factors. The model is estimated on the California add-on data set of the 2009 National Household Travel Survey. Model estimation results show that spatial dependency effects are statistically significant. The findings have important implications for model development and application in the policy forecasting arena.

INTRODUCTION

The contribution of transportation to energy consumption and greenhouse gas emissions is undoubtedly dependent on the nature of vehicular travel undertaken by households. The number of vehicles owned, the types of vehicles owned (in terms of size, weight, fuel type, and age), and the extent to which different vehicles are used (miles of travel) are all key determinants of energy consumption and greenhouse gas emissions. Over the past 25 years, the split between cars and light duty trucks in the nation's vehicle fleet has changed dramatically; whereas light duty trucks (including pick-up trucks, minivans, and sport utility vehicles) accounted for just about 20 percent of the fleet 25 years ago, they now account for about one-half of all vehicles on the nation's roadways (1). This dramatic shift in the vehicular fleet composition and utilization has had far reaching energy and environmental consequences.

The impact of the composition and utilization of the household vehicular fleet on energy consumption and greenhouse gas emissions calls for the incorporation of behavioral models of vehicle type choice and utilization in transportation demand forecasting models. Such models would provide the ability to forecast energy and environmental impacts of shifting vehicle ownership and utilization patterns arising from alternative policy decisions, the advent of new alternative fuel vehicle technologies, and changes in household and personal vehicular preferences. In this context, while there have been several earlier efforts in the literature on vehicle ownership analysis, much remains to be done in developing behavioral models of household vehicle fleet composition and utilization choices – and connecting such choices to energy and emissions estimates.

In particular, an important issue that has not been adequately addressed in the vehicle ownership and utilization literature is that there may be spatial interaction effects in household vehicle ownership and type choice that are both observed and unobserved. Vehicle choices that households make are likely to be influenced by their interactions with neighboring households and the choices that neighboring households make. If a household observes that many of its neighbors own and drive hybrid electric vehicles, or hears good reviews about such vehicles from neighbors who already own and drive them, then the household may be motivated and influenced to also own and drive a hybrid electric vehicle. Spatial interaction effects may also arise from unobserved attitudinal preferences whereby households with similar lifestyle preferences cluster together in neighborhoods that have built environment attributes conducive to their lifestyle choices.

This paper aims to contribute to the vehicle ownership and fleet composition analysis literature by presenting a multinomial probit model that explicitly accounts for spatial interaction effects in these choice phenomena. Underlying the multinomial probit model with spatial interaction effects is a behavioral framework that not only estimates the number of vehicles owned by a household, but also the vehicle type choice – thus allowing the construction of the entire vehicle fleet for a household, while explicitly considering spatial dependency effects.

SPATIAL DEPENDENCE IN CHOICE MODELING

The past decade has seen increasing attention being paid to accommodating spatial dependency effects in modeling choice-making behaviors of agents in a variety of contexts (2). There have been several efforts in the recent past to apply spatial correlation structures that have been developed for modeling continuous dependent variables in the context of discrete choice models of behavior (see recent reviews of this literature in Anselin (2) and Bhat *et al.* (3)). However, these efforts have been hampered by the need to evaluate multidimensional integrals of the order

of the product of the number of decision agents and the number of alternatives minus one for unordered multinomial response choice models.

Several studies (4,5) have side-stepped the high-dimensional problem inherent in global and general spatial dependency structures by assuming that the dependency originates only from observed exogenous covariates of proximate decision agents. However, this is rather untenable in the context of several choice situations where the spatial dependence naturally arises from didactic interactions between decision agents. To elucidate, households may be viewed as developing utilities (or preferences) for vehicle type choice alternatives based on a set of observed factors (such as income and presence of children in neighboring households) as well as unobserved tastes, attitudes, and location factors (such as how “green” a household is in its views and whether there are continuous sidewalks/bicycle paths in the neighborhood). The utility vector of one household is likely to be influenced by the utility vector of other nearby households due to didactic interactions and interchanges (where utility signals get bounced around across decision agents). In this process, there is a “spatial spillage” effect not only based on the observed covariate effects of neighboring households, but also due to unobserved factors. For example, a neighboring household’s perception of “greenness” or the quality of sidewalks/bike paths may spill over and influence choices of another household. Further, there may be residential self-selection effects leading to a sorting of households based on similarity in unobserved vehicle type choice preferences.

In discrete choice models, ignoring these spillage effects due to observed factors and/or due to unobserved factors will, in general, lead to inconsistent estimates of the effects of observed covariates. As indicated by Anselin (6), it behooves the analyst to include spatial “spillover” effects in both the observed covariates as well as the errors unless there are strong a priori reasons not to do so. In the current paper, a spatial lag formulation is adopted to accommodate global spatial dependence effects (due to both observed covariate and error spillage effects) in household vehicle type choice decisions. The specific model structure and formulation implemented in this paper allows the modeling of the entire vehicle fleet composition of households, as is discussed in the next section. The development of a multinomial probit model with continuous spatial dependency effects (due to both observed and unobserved factors) that is capable of modeling the entire vehicle fleet composition constitutes the novel contribution of this paper.

DATA

The data set used in this study is derived from the California add-on component of the 2009 National Household Travel Survey (NHTS). The National Household Travel Survey (NHTS) is a national survey conducted by the United States Department of Transportation to measure the amount of personal travel that is undertaken by the nation’s populace. Individual states and metropolitan areas are allowed to purchase and commission additional data collection within their jurisdictions if they desire larger samples for their own analysis and planning applications. Within the California add-on survey sample, the subsample from the Los Angeles city region was extracted for the analysis conducted in this paper. As spatial interaction effects are likely to be more localized in nature, it was considered prudent to use a data set from a limited geographic region. The desire to limit the sample size (and thus avoid inflated t-statistics that might arise from the use of large samples) was another consideration in the selection of a subsample from a limited geographic region. Finally, the selection of this specific subsample made it possible to merge census tract level accessibility measures and land use data that have been compiled in

connection with an ongoing parallel effort to develop a comprehensive activity-based microsimulation model system for the Southern California Association of Governments (7). The accessibility measures are opportunity-based indicators which measure the number of activity opportunities by 12 different industry types as well as total roadway length of different roadway types that can be reached within 10 minutes using the auto mode from the home census tract during the morning peak period (6 AM to 9 AM).

The data set includes detailed individual and household level socio-economic and demographic data together with information about the vehicle fleet in each household. After extensive cleaning and filtering for missing data, a survey sample of 961 households was available for analysis. In order to limit the sample size and for reasons of computational tractability, a 25 percent random sample of 243 households residing in 200 census tracts was chosen for model estimation. For the model estimation exercise in this paper, vehicle type choice was represented as a combination of two dimensions – body type and vintage. Two body types were considered, namely, car and non-car (encompassing sport utility vehicles, vans/minivans, and pick-up trucks). Two age categories were considered – less than or equal to five years old, and greater than five years old. Thus there are four vehicle type alternatives defined in this paper.

An examination of the descriptive characteristics of the sample of 243 households suggests that the data set is suitable for the model estimation effort undertaken in this paper. It is found that 8.2 percent of households have no vehicle, another 34.5 percent have one vehicle, and 40 percent have two vehicles. Among the vehicles in the sample, 40 percent are old cars, 24 percent are new cars (less than or equal to five years old), another 24 percent are old non-cars, and 12 percent are new non-cars. Among other descriptive statistics, 82 percent of the households are of non-Hispanic origin, with 68 percent of individuals reporting their race as Caucasian. About 70 percent of households own the home in which they reside. With respect to the income distribution, it is found that one-fifth of the households report an annual income less than \$20,000 and an equal proportion report incomes between \$20,000 and \$45,000. Just about 38 percent of the households report income greater than \$75,000 per year. About 47 percent of the households report having one adult and another 46 percent report having two adults. Nearly 34 percent of households have zero workers, and 44 percent have one worker. About 17 percent of households report having one self-employed individual. There is one person with more than one job in 11 percent of the households. The employed individuals report a mean distance to work of 6.1 miles. Only one percent of the households report having a child 0-5 years of age, but 12 percent of households report having a child 6-10 years of age. About 12 percent of households report having a child 11-15 years of age (households not necessarily mutually exclusive). Just over one-third of households report having a senior adult who is 65 years of age or older. About 35 percent of households are immigrant households. The mean distance between households (based on the census tracts of household residences), which is the distance measure used to capture spatial dependence effects due to proximity, is 11.1 miles with a standard deviation of 6.6 miles. The corresponding median distance is 11 miles. Additionally, 20.4% of household pairings have inter-household distances of less than 5 miles. Thus, there are enough households close to one another, as well as enough variation in the inter-household distances across household pairings, to estimate spatial dependency effects.

MODELING METHODOLOGY

The behavioral framework adopted in this study assumes that the observed vehicle fleet of a household is the result of a series of unobserved (to the analyst) repeated “synthetic” discrete choice occasions in which the household chooses not to purchase a vehicle or chooses a vehicle of a certain type. The number of synthetic choice occasions in such a “vertical” (over time) choice setting is linked to the number of driving age members in the household to exploit the fact that the number of vehicles owned by a household is virtually never greater than the number of driving age members (say N) plus two (in the data set used in the current analysis, 99.1% of households were covered by this condition). Thus, for each household, a set of $N+2$ synthetic choice occasions is created and an appropriate choice is assigned as the dependent variable. For estimation, there needs to be a procedure to assign a chosen alternative at each synthetic occasion. For this, the temporal sequence of vehicle purchases of the household, as reported in the survey, is used. For example, say a household owns an old sedan and a new sports utility vehicle (SUV), with the old sedan being purchased first. Then, the old sedan is the chosen alternative at the first choice occasion, and the new SUV is the chosen alternative in the second. The chosen alternative in the remaining two choice occasions is “no vehicle purchased”. For the second choice occasion, information that the household already has an old sedan is used as an explanatory variable.¹ The procedure above mimics the dynamics of fleet ownership decisions, although there is no temporal component of the dynamics involved because only synthetic choice occasions are considered; the observed information available is only that of vehicles held at a cross-sectional point in time with information on the sequence in which the currently held vehicles were purchased.²

Modeling Approach

Let the instantaneous utility U_{qti} of household q ($q = 1, 2, \dots, Q$) at synthetic choice occasion t ($t = 1, 2, \dots, T_q$) for vehicle type choice i ($i = 1, 2, \dots, I$; $I = 5$ in the empirical context of the current paper, including the “no vehicle purchase” alternative) be a function of a $(K \times 1)$ -column vector of exogenous attributes \mathbf{x}_{qti} (including household demographics, types of vehicles “chosen” before the t^{th} choice occasion, and activity-travel environment characteristics). Let $T_q = N_q + 2$, where T_q is the number of synthetic choice occasions for household q , and N_q is the number of driving age members in household q . Note that t does not have a chronological time interpretation. It is simply a device to accommodate multiple synthetic choice occasions and mimic the dynamics of fleet ownership decisions. That is, $t=1$ for a household A does not have any chronological time bearing to $t=1$ or $t=2$ for a neighboring household B. However, the choice occasions of different households may be considered to occur over a time period wherein households are interacting and exchanging utility signals. Thus, the spatial dependence across households is specified for each vehicle type i without any specific association to the choice occasion. That is, the utility

¹ It is also possible to assign the old sedan to the first choice occasion, no vehicle in the second, no vehicle in the third, and the new SUV in the fourth occasion. However, both of these assignments give the same results, because the “dynamics” are based on what the household already owns in total, not what was chosen in the immediately previous choice occasion.

² To be sure, our approach is not a true vehicle fleet evolution model that analyzes the dynamics of vehicle transaction decisions over time. The estimation of such evolution models, while appealing from a behavioral standpoint, has been hampered by the paucity of longitudinal data on vehicle transactions. Moreover, many dynamic models have focused primarily on vehicle ownership (*i.e.*, transactions) with inadequate emphasis on the vehicle type, usage, and vintage considerations of the household fleet.

U_{qti} for household q at choice occasion t for alternative i is related to the utility $U_{q't'i}$ for household q' and alternative i at each (and all) of the choice occasions t' ($t'=1, 2, \dots, T_{q'}$) of household q' . This is an important distinction from the traditional spatial dependency specifications for spatial panel discrete choice models, and leads to a specific form for the model in this study that has not appeared previously in the literature.

Thus, the utility U_{qti} incorporating a spatial lag structure is written as follows:

$$U_{qti} = \delta \sum_{q'} w_{qq'} \sum_{t'=1}^{T_{q'}} U_{q't'i} + \tilde{\alpha}_{qi} + \boldsymbol{\beta}'_q \mathbf{x}_{qti} + \tilde{\varepsilon}_{qti} \quad (1)$$

where $w_{qq'}$ is a distance-based spatial weight corresponding to units q and q' (with $w_{qq} = 0$ and $\sum_{q'} w_{qq'} = 1$) for each (and all) q , δ ($0 < \delta < 1$) is the spatial lag autoregressive parameter, $\tilde{\alpha}_{qi}$ is a normal random-effect term capturing a household-specific stationary preference effect for vehicle type i , and $\boldsymbol{\beta}_q$ is a household-specific ($K \times 1$)-vector of coefficients assumed to be a realization from a multivariate normal distribution with mean vector \mathbf{b} and covariance $\tilde{\boldsymbol{\Omega}} = \mathbf{L}\mathbf{L}'$. Let $\boldsymbol{\beta}_q = \mathbf{b} + \tilde{\boldsymbol{\beta}}_q$, where $\tilde{\boldsymbol{\beta}}_q \sim MVN_K(0, \tilde{\boldsymbol{\Omega}})$ is a mixing (multivariate) distribution to capture unobserved sensitivity variations (to the exogenous variables in the vector \mathbf{x}_{qti}) across households (MVN_K represents the multivariate normal distribution of dimension K). Also, write $\tilde{\alpha}_{qi} = \tilde{\alpha}_i + \tilde{\alpha}_{qi}$, and let the mean and variance-covariance matrix of the vertically stacked ($I \times 1$)-vector of random-effect terms $\tilde{\boldsymbol{\alpha}}_q \left[= (\tilde{\alpha}_{q1}, \tilde{\alpha}_{q2}, \dots, \tilde{\alpha}_{qI})' \right]$ be $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{\Lambda}}$, respectively. $\tilde{\varepsilon}_{qti}$ in Equation (1) is a normal error term uncorrelated with $\tilde{\boldsymbol{\beta}}_q$ and all $\tilde{\alpha}_{qi}$ terms ($i = 1, 2, \dots, I$), and also uncorrelated across observation units q and synthetic choice occasions t . However, at each synthetic choice occasion t for household q , the $\tilde{\varepsilon}_{qti}$ terms may have a covariance (dependency) structure across vehicle types i due to choice-occasion unobserved factors that simultaneously increase or simultaneously decrease the utility of certain types of vehicles: $\tilde{\boldsymbol{\varepsilon}}_{qt} \left[= (\tilde{\varepsilon}_{qt1}, \tilde{\varepsilon}_{qt2}, \dots, \tilde{\varepsilon}_{qtI})' \right] \sim MVN_I(0, \tilde{\boldsymbol{\Psi}})$.

As usual, appropriate scale and level normalization must be imposed on $\tilde{\mathbf{A}}$, $\tilde{\mathbf{\Lambda}}$ and $\tilde{\boldsymbol{\Psi}}$ for identification purposes. Specifically, only utility differentials matter in discrete choice models. At the same time, whenever utility differentials are taken during estimation, they must all originate from the same underlying matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{\Lambda}}$ and $\tilde{\boldsymbol{\Psi}}$. To achieve this, take the utility differentials with respect to the first alternative. Then, only the elements $\alpha_{q1} = \tilde{\alpha}_{qi} - \tilde{\alpha}_{q1}$ ($i \neq 1$) and its covariance matrix $\boldsymbol{\Lambda}_1$, and the covariance matrix $\boldsymbol{\Psi}_1$ of $\xi_{qti} = \tilde{\varepsilon}_{qti} - \tilde{\varepsilon}_{q1i}$ ($i \neq 1$), are estimable. So, a normalization $\tilde{\alpha}_{q1} = 0 \forall q$ is applied, implying that $\tilde{\alpha}_1 = 0$. Also, develop $\boldsymbol{\Lambda}$ from $\boldsymbol{\Lambda}_1$ by adding an additional row on top and an additional column to the left. All elements of this additional row and additional column are filled with values of zeros. Similarly, construct $\boldsymbol{\Psi}$ from $\boldsymbol{\Psi}_1$ by adding a row on top and a column to the left. This first row and the first column of the matrix $\tilde{\boldsymbol{\Psi}}$ are also filled with zero values. An additional normalization needs to be imposed

on $\tilde{\Psi}$ because the scale is also not identified. For this, normalize the element of $\tilde{\Psi}$ in the second row and second column to the value of one. Note that all of these normalizations are needed for econometric identification purposes.

Next, define the following:

$$\mathbf{U}_{qt} = (U_{qt1}, U_{qt2}, \dots, U_{qtI})', \quad \tilde{\boldsymbol{\varepsilon}}_{qt} = (\tilde{\boldsymbol{\varepsilon}}_{qt1}, \tilde{\boldsymbol{\varepsilon}}_{qt2}, \dots, \tilde{\boldsymbol{\varepsilon}}_{qtI})' \quad (I \times 1 \text{ vectors}),$$

$$\mathbf{U}_q = (\mathbf{U}'_{q1}, \mathbf{U}'_{q2}, \dots, \mathbf{U}'_{qT_q})', \quad \tilde{\boldsymbol{\varepsilon}}_q = (\tilde{\boldsymbol{\varepsilon}}'_{q1}, \tilde{\boldsymbol{\varepsilon}}'_{q2}, \dots, \tilde{\boldsymbol{\varepsilon}}'_{qT_q})' \quad ((T_q \times I) \times 1 \text{ vectors}),$$

$$\mathbf{U} = (\mathbf{U}'_1, \mathbf{U}'_2, \dots, \mathbf{U}'_Q)', \quad \tilde{\boldsymbol{\varepsilon}} = (\tilde{\boldsymbol{\varepsilon}}'_1, \tilde{\boldsymbol{\varepsilon}}'_2, \dots, \tilde{\boldsymbol{\varepsilon}}'_Q)' \quad (RI \times 1 \text{ vectors}), \quad R = \sum_{q=1}^Q T_q,$$

$$\tilde{\boldsymbol{\alpha}}_q = (\tilde{\alpha}_{q1}, \tilde{\alpha}_{q2}, \dots, \tilde{\alpha}_{qI})' \quad (I \times 1 \text{ vector}), \quad \tilde{\boldsymbol{\alpha}} = [(1_{T_1} \otimes \tilde{\alpha}_1)', (1_{T_2} \otimes \tilde{\alpha}_2)', \dots, (1_{T_Q} \otimes \tilde{\alpha}_Q)']' \quad (RI \times 1 \text{ vector}),$$

$$\mathbf{x}_{qt} = (\mathbf{x}_{qt1}, \mathbf{x}_{qt2}, \dots, \mathbf{x}_{qtI})' \quad (I \times K \text{ matrix}), \quad \mathbf{x}_q = (\mathbf{x}'_{q1}, \mathbf{x}'_{q2}, \dots, \mathbf{x}'_{qT_q})' \quad ((T_q \times I) \times K \text{ matrix}),$$

$\mathbf{x} = (\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_Q)' \quad (RI \times K \text{ matrix}),$ and $\tilde{\boldsymbol{\beta}} = (\tilde{\boldsymbol{\beta}}'_1, \tilde{\boldsymbol{\beta}}'_2, \dots, \tilde{\boldsymbol{\beta}}'_Q)' \quad (QK \times 1 \text{ vector})$. Let \mathbf{IDEN}_E be the identity matrix of size E , $\mathbf{1}_E$ be a column vector of size E with all of its elements taking the value of one, and $\mathbf{1}_{EE}$ be a square matrix of size E with all unit elements. Also, define the following matrix:

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_1 & 0 & 0 & 0 \dots 0 \\ 0 & \mathbf{x}_2 & 0 & 0 \dots 0 \\ 0 & 0 & \mathbf{x}_3 & 0 \dots 0 \\ \vdots & \vdots & \vdots & \dots \vdots \\ 0 & 0 & 0 & 0 \dots \mathbf{x}_Q \end{bmatrix} \quad (RI \times QK \text{ matrix}), \quad (2)$$

Let $\tilde{R}_q = \sum_{j=1}^{q-1} T_j$, with the convention that $\tilde{R}_1 = 0$, and let $\tilde{R}_q = \tilde{R}_q \times I$. Define a matrix \mathbf{C} of size

$RI \times RI$ that is filled with sub-matrices of size $(T_q \times I) \times (T_{q'} \times I)$ as follows: $[\mathbf{C}]_{\{(\tilde{R}_q+1)-\tilde{R}_{q+1}\}, \{(\tilde{R}_{q'}+1)-\tilde{R}_{q'+1}\}} = w_{qq'} \otimes \mathbf{1}_{T_q T_{q'}} \otimes \mathbf{IDEN}_I$, where $[\mathbf{C}]_{\{(\tilde{R}_q+1)-\tilde{R}_{q+1}\}, \{(\tilde{R}_{q'}+1)-\tilde{R}_{q'+1}\}}$ refers to the sub-matrix of \mathbf{C} that corresponds to the $(\tilde{R}_q + 1)^{th}$ through \tilde{R}_{q+1}^{th} rows and $(\tilde{R}_{q'} + 1)^{th}$ through $\tilde{R}_{q'+1}^{th}$ columns.

Let $\mathbf{S} = [\mathbf{IDEN}_{RI} - \delta \mathbf{C}]^{-1}$ ($RI \times RI$ matrix) and

$\tilde{\mathbf{A}} = [(1_{T_1} \otimes \tilde{\mathbf{A}})', (1_{T_2} \otimes \tilde{\mathbf{A}})', \dots, (1_{T_Q} \otimes \tilde{\mathbf{A}})']' \quad (RI \times 1 \text{ matrix})$. Then, Equation (1) may be written in matrix notation as:

$$\mathbf{U} = \mathbf{S}[\tilde{\mathbf{A}} + \mathbf{x}\mathbf{b} + \tilde{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}}]. \quad (3)$$

Let $[\cdot]_e$ indicate the e^{th} element of the column vector $[\cdot]$, and let $d_{qti} = \tilde{R}_q + (t-1)I + i$. Equation (3) can be equivalently written as:

$$U_{qti} = [\mathbf{S}\{\tilde{\mathbf{A}} + \mathbf{x}\mathbf{b}\}]_{d_{qti}} + [\mathbf{S}\{\tilde{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}}\}]_{d_{qti}}. \quad (4)$$

Define $V_{qti} = [\mathbf{S}\{\tilde{\mathbf{A}} + \mathbf{x}\mathbf{b}\}]_{d_{qti}}$ and $\varepsilon_{qti} = [\mathbf{S}\{\tilde{\boldsymbol{\alpha}} + \tilde{\mathbf{x}}\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\varepsilon}}\}]_{d_{qti}}$. Household q chooses the vehicle type at synthetic choice occasion t that provides maximum utility. Let the ‘‘chosen’’ vehicle type

(assigned as described previously) for household q at occasion t be m_{qt} . In the utility differential form, Equation (4) may be written as:

$$y_{qtim_{qt}} = U_{qti} - U_{qtm_{qt}} = H_{qtim_{qt}} + \xi_{qtim_{qt}}; H_{qtim_{qt}} = V_{qti} - V_{qtm_{qt}} \text{ and } \xi_{qtim_{qt}} = \varepsilon_{qti} - \varepsilon_{qtm_{qt}}; i \neq m_{qt} \quad (5)$$

Then stack the utility differentials $y_{qtim_{qt}} (= U_{qti} - U_{qtm_{qt}}, i \neq m_{qt})$ in the following order: $\mathbf{y}_{qt} = (y_{qt1m_{qt}}, y_{qt2m_{qt}}, \dots, y_{qtlm_{qt}})'$, an $(I-1) \times 1$ vector; $\mathbf{y}_q = (\mathbf{y}'_{q1}, \mathbf{y}'_{q2}, \dots, \mathbf{y}'_{qT_q})'$, an $[(I-1) \times T_q] \times 1$ vector; and $\mathbf{y} = (\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_Q)'$, an $[(I-1) \times R] \times 1$ vector. Correspondingly, let $\mathbf{H}_{qt} = (H_{qt1m_{qt}}, H_{qt2m_{qt}}, \dots, H_{qtlm_{qt}})'$, an $(I-1) \times 1$ vector; $\mathbf{H}_q = (\mathbf{H}'_{q1}, \mathbf{H}'_{q2}, \dots, \mathbf{H}'_{qT_q})'$, an $[(I-1) \times T_q] \times 1$ vector; and $\mathbf{H} = (\mathbf{H}'_1, \mathbf{H}'_2, \dots, \mathbf{H}'_Q)'$, an $[(I-1) \times R] \times 1$ vector. It may be noted that \mathbf{y} has a mean vector \mathbf{H} .

To determine the covariance matrix of \mathbf{y} , several additional matrix definitions are needed. Define a matrix $\mathbf{\Lambda}$ of size $RI \times RI$ that is block-diagonal, with each block diagonal as follows: $[\mathbf{\Lambda}]_{\{(\bar{R}_q+1)-\bar{R}_{q+1}\}, \{(\bar{R}_q+1)-\bar{R}_{q+1}\}} = \mathbf{1}_{T_q T_q} \otimes \tilde{\mathbf{\Lambda}} \quad (q=1, 2, \dots, Q)$, $\mathbf{\Omega} = \tilde{\mathbf{x}}(\mathbf{I}_Q \otimes \tilde{\mathbf{\Omega}})\tilde{\mathbf{x}}'$ ($RI \times RI$ matrix), and $\mathbf{\Psi} = \mathbf{IDEN}_R \otimes \tilde{\mathbf{\Psi}}$ ($RI \times RI$ matrix). Let $\tilde{\mathbf{F}} = \mathbf{S}[\mathbf{\Lambda} + \mathbf{\Omega} + \mathbf{\Psi}]\mathbf{S}'$ and define \mathbf{M} as an $(q = 1, 2, \dots, Q)$: $[(I-1) \times R] \times [I \times R]$ block diagonal matrix, with each block diagonal having $(I-1)$ rows and I columns corresponding to the t^{th} synthetic choice occasion of household q . This $(I-1) \times I$ matrix for household q and choice occasion t corresponds to an $(I-1)$ identity matrix with an extra column of -1 's added as the m_{qt}^{th} column. Finally, the multivariate distribution of the utility differentials is obtained, $\mathbf{y} : \mathbf{y} \sim MVN(\mathbf{H}, \mathbf{\Sigma})$, where $\mathbf{\Sigma} = \mathbf{M}\tilde{\mathbf{F}}\mathbf{M}'$. Next, let $\boldsymbol{\theta}$ be the collection of parameters to be estimated: $\boldsymbol{\theta} = [\mathbf{b}'; \text{Vech}(\tilde{\mathbf{\Omega}}); \tilde{\mathbf{\Lambda}}', \text{Vech}(\tilde{\mathbf{\Lambda}}), \text{Vech}(\tilde{\mathbf{\Psi}}), \boldsymbol{\delta}]'$, where $\text{Vech}(\tilde{\mathbf{\Omega}})$ represents the row vector of upper triangle elements of $\tilde{\mathbf{\Omega}}$. Then, the likelihood of the observed sample may be written succinctly as $\text{Prob}[\mathbf{y}^* < \mathbf{0}]$.

$$L_{ML}(\boldsymbol{\theta}) = \text{Prob}[\mathbf{y}^* < \mathbf{0}] = F_{R \times (I-1)}(-\mathbf{H}, \mathbf{\Sigma}) \quad (6)$$

where $F_{R \times (I-1)}$ is the multivariate cumulative normal distribution of $R \times (I-1)$ dimensions. Despite advances in simulation techniques and computational power, the evaluation of such a high dimensional integral is literally infeasible using established estimation techniques.

Model Estimation Procedure

In view of the computational intractability of the likelihood function presented earlier, the current study uses Bhat's (8) maximum approximate composite marginal likelihood (MACML) inference approach in estimation. The MACML approach combines a composite marginal likelihood (CML) estimation approach with an approximation method to evaluate the multivariate standard normal cumulative distribution (MVNCD) function. The CML approach works as follows. Instead of developing the likelihood of the entire sample, consider developing a surrogate likelihood function that is the product of the probability of easily computed marginal events. For instance, one may compound (multiply) pairwise probabilities of household q choosing the actual "chosen" vehicle type m_{qt} at occasion t and choosing the actual "chosen" vehicle type m_{qs} at occasion s , of household q choosing vehicle type m_{qt} at occasion t and household q' choosing vehicle type $m_{q's}$ at time s , and so on. The CML estimator is then the one

that maximizes the compounded probability of all pairwise events. The CML function may be written as:

$$L_{CML}(\boldsymbol{\theta}) = \prod_{q=1}^Q \prod_{q'=q}^Q \prod_{t=1}^T \prod_{t'=t}^T \text{Prob}(C_{qt} = m_{qt}, C_{q't'} = m_{q't'}) \text{ with } q \neq q' \text{ when } t = t', \quad (7)$$

where C_{qt} is an index for the vehicle type chosen by household q at occasion t . Each of these pairwise probabilities is of $(I-1) \times 2$ dimensions, which may be computed easily using the MVNCD approximation method embedded in the MACML method. The pairwise marginal likelihood function of Equation (7) comprises $R(R-1)/2$ pairs of multivariate pairwise probability computations, which can itself become quite time consuming. Fortunately, in a spatial-temporal case where spatial dependency drops quickly with inter-observation distance, it should suffice to retain pairs within a certain threshold distance. This threshold value is estimated by testing different distance bands, starting from a small distance band and increasing the band. Then, the asymptotic variance matrix $V_{CML}(\hat{\boldsymbol{\theta}})$ is estimated for each distance band and the threshold distance value (say \tilde{d}_{thresh}) is chosen as the value beyond which there is either an increase or no additional decrease in the total variance across all parameters as given by $tr[V_{CML}(\hat{\boldsymbol{\theta}})]$ (i.e., the trace of the matrix $[V_{CML}(\hat{\boldsymbol{\theta}})]$).

The CML estimator of $\boldsymbol{\theta}$ is consistent and asymptotically normal distributed with asymptotic mean $\boldsymbol{\theta}$ and covariance matrix given by the inverse of Godambe's sandwich information matrix (9). Bhat (8) provides details of how to compute the covariance matrix.

In the estimations, the positive-definiteness of each of the $\tilde{\boldsymbol{\Omega}}$, $\tilde{\boldsymbol{\Lambda}}$, and $\tilde{\boldsymbol{\Psi}}$ matrices is guaranteed by writing the logarithm of the pairwise-likelihood in terms of the Cholesky-decomposed elements of these matrices, and maximizing with respect to these elements of the Cholesky factor. To ensure the constraint $0 < \delta < 1$, this term is parameterized as $\delta = 1/[1 + \exp(\tilde{\delta})]$. Once estimated, the $\tilde{\delta}$ estimate can be translated back to obtain estimates of δ .

ESTIMATION RESULTS

This section presents results of the estimation of the multinomial probit model with spatial dependency effects on the California add-on data set of the 2009 National Household Travel Survey. Estimation results are presented in Table 1. A number of model specifications were estimated prior to arriving at the final model specification. None of the mixing parameters came out to be statistically significant in the final model specification. This result indicates that there is no significant household-specific heterogeneity in the variable effects on the vehicle type choice decisions. It is important to note that, even in the absence of mixing on variables, the model does not collapse to a cross-sectional spatial model. This is because the setup of the model is such that the utility associated with an alternative for one household at any given synthetic choice occasion is influenced by the utility associated with the same alternative across all synthetic choice occasions of all other households in the region, thus leading to a pseudo unbalanced panel setup due to unequal number of choice occasions across individuals. Another key finding is that there were no significant deviations in the error covariance matrix $\boldsymbol{\Psi}_1$ of $\xi_{qit1} = \tilde{\varepsilon}_{qit} - \tilde{\varepsilon}_{q1}$ ($i \neq 1$) from

the corresponding matrix in an independent multinomial probit (MNP) model.³ This finding implies that at any given synthetic choice occasion the choice occasion specific unobserved factors that influence the utility associated with different vehicle type alternatives are all independent and identically distributed.

Although the magnitudes of the constants cannot be directly compared across different alternatives (because there are continuous variables in the utility formulations), the relative values may be loosely interpreted as providing an indication of the baseline preference for different vehicle types. It appears that new cars are the least preferred vehicle type while old cars are the most preferred; there is little difference in the baseline preference between old and new non-cars. It is found that households with higher levels of educational attainment are less inclined to acquire new non-cars. It is possible that these households are more environmentally conscious and savvy consumers shunning the expense and environmental consequences of driving new non-cars (sport utility vehicles, trucks, and vans). Those of Hispanic origin show a greater inclination to acquire old non-cars, while African Americans are less likely to acquire old cars and new non-cars.

Households that own their home have a greater utility for all vehicle alternatives (in comparison to households that do not own their home), with a particular preference for new vehicles in comparison to older vehicles. As expected households in the highest income category have a positive utility for all vehicle alternatives, with a higher preference for new non-cars and the lowest preference for old non-cars. However, the difference in magnitudes of coefficients across vehicle type alternatives is quite modest. As the number of adults increases, households are more likely to acquire old cars, new cars, or old non-cars – presumably these households have a greater need for multiple cars and show a greater disinclination to acquiring new non-cars in view of budget constraints. However, with the easing of budget constraints that invariably comes with greater number of workers in the household, it is found that households show a greater preference towards new cars or non-cars and shun older cars.

As the mean distance to work increases, households are more likely to acquire new non-cars presumably because people are looking for larger and more comfortable and reliable vehicles for the longer commute. The presence of children is associated with a smaller likelihood of acquiring older cars, and a greater likelihood of acquiring older non-cars. Presumably, such households prefer larger cars for the space, and newer cars for the reliability factor. Immigrant households have a lower utility across all vehicle type alternatives compared to non-immigrant households, but have a smaller negative coefficient on the non-car alternatives. Immigrant households may be located in more dense neighborhoods and may be more walk and transit oriented, thus contributing to the lower utility across all vehicle types. The coefficients on the non-car alternatives are less negative, probably because these households are of larger size motivating the acquisition of non-cars in preference to cars. Households with senior adults are less likely to acquire non-cars. There may be two explanations for this; first, seniors may have diminishing driving skills that make the driving and control of larger vehicles cumbersome, and second, seniors may be living in smaller households (empty nests) and so do not need larger vehicles (non-cars) any more.

As the temporal sequence in which vehicles were acquired in the household is known in the survey data set, information about the existing vehicle fleet was used as explanatory variables in the utility specification of vehicle type alternatives for all choice occasions after the first. This

³ The covariance matrix of the error term differences in an independent MNP model has 1 and 0.5 as the diagonal and off-diagonal elements, respectively.

specification mimics the underlying dynamics in the purchase decisions with existing vehicles in the household influencing the vehicles that households acquire subsequently. As expected, parameter estimates for all vehicle types are negative suggesting that households increasingly choose to acquire “no vehicle” as they build up their vehicle fleet. The relative magnitudes of the coefficients can be used to draw inferences about how households tend to construct their fleets. It appears that households are somewhat variety seeking; for example, as the number of old cars increases, households are more prone to acquire new non-cars (least negative coefficient); as the number of new cars increases, households are more prone to acquire new non-cars (and shun older cars); as the number of old non-cars increases, households are more prone to acquire new non-cars or new cars; finally, as the number of new non-cars increases, households are more prone to acquire new cars. In all cases, the least negative coefficient is associated with a car type different than that representing the explanatory variable for the existing vehicle fleet.

With respect to accessibility measures, households with good access to primary and minor arterials have a lower preference for older non-cars suggesting that these households may be more auto-oriented (and hence located in census tracts with good roadway presence) and prefer to drive newer cars. Households in census tracts closer to manufacturing employment are less likely to acquire new cars, possibly because these census tracts are in lower income areas thus making new car purchases challenging. On the other hand, households in census tracts close to “arts” employment are less likely to acquire old cars; it is possible that these census tracts are in urban arts districts that are trendier and people who locate there are more likely to acquire newer cars.

MODEL ASSESSMENT

The model estimation effort yielded coefficient values that are largely reasonable and behaviorally intuitive. This section offers an assessment of the model from a number of different perspectives including the significance of the spatial dependency parameter, the goodness of fit of the model relative to a model that does not include spatial dependence, and differences in elasticity estimates between the multinomial probit model with spatial dependency and the independent multinomial probit model that ignores spatial dependency.

Among the different weight matrix specifications that were tested, the inverse distance based specification was found to offer the best fit. The spatial autoregressive parameter in the spatial lag formulation, δ , also turns out to be statistically significant with a value of 0.1872 and t-statistic of 3.80. This is evidence of the presence of spatial spillover effects arising either due to didactic interactions of individuals in proximally located households or due to residential self-selection effects that can lead to a clustering of households with similar vehicle type choice preferences.

Although the spatial parameter is statistically significant suggesting superior data fit in the spatial model compared to a corresponding non-spatial model, an alternative way to compare these nested models is through the adjusted composite likelihood ratio (ADCLRT) test (8). The composite log-likelihood value for the non-spatial model is -138971.8 (52 parameters estimated) and that for the final spatial model is -138827.6 (53 parameters estimated). The ADCLRT test statistic of comparison between the two models is 7.84 which is greater than the critical chi-squared value of 3.84 associated with one degree of freedom, thus demonstrating presence of spatial interactions in vehicle type choice decisions. Ignoring such spatial interaction effects results in a model with poorer statistical goodness-of-fit.

A question that is often raised in the context of advanced choice models that incorporate additional (observed or unobserved) effects is the extent to which policy forecasts might actually differ due to ignoring such effects. Although the goodness-of-fit is significantly better and the spatial interactions parameter is significant, does that mean that policy forecasts would be different as a result of using one model versus the other (that ignores spatial dependence effects)? The parameter estimates in Table 1 do not directly provide the magnitude of the impact of variables on the probabilities of acquiring each vehicle type alternative. To shed light on this question, it is useful to compute aggregate-level elasticity effects of variables for the different model specifications. Specifically, the effects of variables on the expected share of each vehicle type alternative are examined in this paper. This is achieved by computing the marginal probability of each household choosing a certain vehicle type in a single synthetic choice scenario and aggregating these probabilities across households and all choice occasions for each vehicle type alternative.

The following procedure is used to compute the shares of each vehicle type alternative. The utility function of vehicle type i for household q is as follows:

$$U_{qii} = \delta \sum_{q'} w_{qq'} \sum_{i'=1}^{T_{q'}} U_{q'i} + \mathbf{b}'\mathbf{x}_{qii} + \tilde{\varepsilon}_{qii} \quad (8)$$

where the notation is similar to that described in the methodology section of this paper. Then, using other notations described previously, it is possible to write:

$$\mathbf{U} = \mathbf{S}[\mathbf{xb} + \tilde{\boldsymbol{\varepsilon}}], \quad (9)$$

The above $RI \times 1$ -vector \mathbf{U} is simulated 500 times using the estimated values of \mathbf{b} , and by randomly drawing 500 times from the appropriate normal distributions for $\tilde{\boldsymbol{\varepsilon}}$. Next, the chosen alternative is determined as the alternative with the highest utility for each of the 500 draws. Finally, the predicted share of each alternative across the 500 draws is taken as an estimate of the probability of each vehicle type alternative. The aggregate share (across all households and all synthetic choice occasions) of each vehicle type alternative is obtained by aggregating the synthetic choice occasion level probabilities of each vehicle type alternative across all households.

The elasticity computed is a measure of the percent change in the aggregate share of each vehicle type alternative due to a change in an exogenous variable. For dummy variables, the value of the variable is changed to one for the subsample of observations for which the variable takes a value of zero, and to zero for the subsample of observations for which the variable takes a value of one. The shifts in expected aggregate shares in the two subsamples are then added after reversing the sign of the shifts in the second subsample, yielding the effective percent change in the expected shares across all households in the sample due to a change in the dummy variable from 0 to 1. For continuous variables, the value of the variable is increased by 25 percent for each observation and the percent change in the expected shares is computed. For variables which only take integer values (such as number of full time workers), the value is increased by unity.

Elasticity estimates are computed for the non-spatial MNP and the spatial MNP model, and are presented in Table 2. The first entry in Table 2 indicates that, according to the MNP model with no spatial interaction, households with a highest education attainment of Bachelor's degree are 3.43% more likely to not acquire a vehicle at any given choice occasion compared to other households. Other elasticity effects can be interpreted similarly.

All of the elasticity effects are consistent with the parameter estimates in Table 1. Also, the elasticity effects of the spatial and non-spatial models are in the same direction (sign) for all variables. However, the elasticity estimates of the non-spatial MNP model and spatial MNP models are quite different *in magnitude*. In general, the elasticity effects of the spatial model are consistently higher in magnitude than those from the non-spatial model. For instance, the elasticity effect of the number of full time workers for the new non-car alternative is 312% where as the corresponding number according to the non-spatial MNP model is only 80%. Similarly, the spatial model implies that Hispanic households are 198% more likely to obtain an old non-car whereas the non-spatial model implies only 85% higher likelihood for Hispanic households. Although the magnitude of the spatial autoregressive parameter is relatively small, the spatial spillover effect is compounding the elasticity estimates due to the circular reinforcing mechanism whereby a change in the value of a variable for one household changes utilities of vehicle type alternatives for other nearby households, which in turn causes ripple effects in the utility values of the household for which the variable changed in the first place.

In summary, it can be seen that elasticity estimates differ substantially between a model with and a model without spatial dependency effects. These differences can have dramatic implications for policy forecasts which often rely on model parameter estimates to infer the magnitude of change in behavior in response to a change in input conditions. For instance, Assembly Bill 32, the Global Warming Solutions Act, requires that the State of California reduce its statewide carbon emissions to 1990 levels by the year 2020, using a combination of transportation, and land use planning strategies. One of these strategies include smart land use development in which proximal housing development is encouraged for developing efficient land use patterns (10). The underlying idea is that high density neighborhoods are less conducive for auto use and thus reduce green house gas (GHG) emissions. However, according to our results, the proximal housing development strategy can be far more effective when coupled with transportation strategies that encourage alternate fuel vehicle (AFV) purchases. Thus, consider a scenario where the local government provides subsidies to purchase AFVs to households in a new dense housing development. Now, due to social interactions among proximal households, other households in the region are also likely to get similar vehicle types. This spatial spillover effect coupled with the auto-restraining neighborhood built environment characteristics can reduce emissions quite substantially. The true potential of such coordinated land-use and transportation planning strategies can be evaluated only using models that account for spatial dependencies of proximal decision making agents, such as the one developed in this paper.

CONCLUSIONS

This paper presents a multinomial probit model of vehicle ownership by type (fleet composition) that explicitly incorporates spatial interaction effects due to observed and unobserved factors. The model is estimated on the Los Angeles region subsample of the California add-on data set of the 2009 National Household Travel Survey which includes a host of accessibility and land use variables critical to vehicle ownership modeling. Underlying the model is a behavioral framework that considers the household vehicle fleet as being constructed over a series of purchase choice occasions, thus providing the ability to endogenously determine the vehicle fleet size while simultaneously incorporating history dependency in the choice model. In other words, the vehicle type that is acquired at any choice occasion is dependent on the existing vehicle fleet in the household comprising vehicles that were acquired at earlier choice occasions. The model considers five choice alternatives for each occasion – two body types (car and non-car) by two

vintage types (less than or equal to five years old, and greater than five years old) plus the choice of acquiring no vehicle at all. The maximum approximate composite marginal likelihood (MACML) estimation procedure is employed to overcome computational intractability associated with traditional simulation and Bayesian model estimation procedures.

Model estimation results show that a host of individual and household variables, not to mention accessibility and land use variables, significantly impact choice of acquiring different vehicle types. More importantly, in the context of this paper, it is found that the spatial interaction parameter is statistically significant and the model that incorporates spatial spillover effects offers a superior statistical goodness-of-fit compared to a multinomial probit model that does not incorporate spatial dependency effects. It is found that a distance based spatial interaction function offers the best fit, with interaction between households dropping off as the distance between households increases. A comparison of elasticity estimates offered by the spatial effects choice model estimated in this paper against those offered by a model with no spatial effects shows that elasticity estimates differ substantially when spatial effects are incorporated. The elasticity estimates from the spatial effects model are consistently higher in magnitude, suggesting that interaction effects amplify the extent to which households change behavior in response to changes in input conditions. Incorporating spatial effects in models of discrete choice behavior can result in substantially different policy forecasts, with clear implications in the transportation planning and policy arena. Future research efforts in this domain could further explore the use of alternative spatial interaction functions, examine whether the findings hold true in other geographical contexts and data sets, investigate model parameter transferability considerations, and attempt to separate unobserved and observed spatial interaction effects by jointly modeling residential location choice with vehicle ownership and fleet composition choice.

ACKNOWLEDGEMENTS

The authors acknowledge the helpful comments of four anonymous reviewers on an earlier version of the paper. The authors are grateful to Lisa Macias for her help in formatting this document.

REFERENCES

1. EPA. *Sources of Greenhouse Gas Emissions: Transportation Sector Emissions*. United States Environmental Protection Agency, Washington, D.C., 2010.
2. Anselin, L. Thirty years of spatial econometrics. *Papers in Regional Science*, Vol. 89, No. 1, 2010, pp. 3-25.
3. Bhat, C.R., I.N. Sener, and N. Eluru. A flexible spatially dependent discrete choice model: formulation and application to teenagers' weekday recreational activity participation. *Transportation Research Part B*, Vol. 44, No. 8-9, pp. 903-921.
4. Mohammadian A., M. Haider, and P.S. Kanaroglou. Incorporating spatial dependencies in random parameter discrete choice models. Presented at the 84th Annual Meeting of the *Transportation Research Board*, Washington D.C., 2005.
5. Adjemian, M.K., C. Lin, and J. Williams. Estimating spatial interdependence in automobile type choice with survey data. *Transportation Research Part A*, Vol. 44, No. 9, 2010, pp. 661-675.
6. Anselin, L. Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review*, Vol. 26, No. 2, 2003, pp. 153-166.
7. Chen, Y., S. Ravulaparthi, L. Deutsch, P. Dalal, S.Y. Yoon, T. Lei, K.G. Goulias, R.M. Pendyala, C.R. Bhat, H.H. Hu. Development of indicators of opportunity-based accessibility. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2255, 2011, pp. 58-68.
8. 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.
9. Godambe, V.P. An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics*, Vol. 31, No. 4, 1960, pp. 1208-1211.
10. California EPA. Climate Action Team Report to Governor Schwarzenegger and the Legislature. California Climate Action Team, 2006.

TABLE 1 Spatial Vehicle Type Choice Model Estimation Results

Variables	Old Car		New Car		Old Non-car		New Non-car	
	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Constant	-0.0351	-0.19	-0.6960	-3.64	-0.5647	-1.92	-0.5432	-1.94
Demographics								
<i>Highest Education Attainment in Household (Base is College degree or less)</i>								
Bachelor Degree	--	--	--	--	--	--	-0.5185	-9.12
Post Graduate	--	--	--	--	--	--	-0.4172	-6.48
<i>Hispanic Status (Base category is non-Hispanic)</i>								
Hispanic Origin	--	--	--	--	0.4983	5.81	--	--
<i>Race (Base category is all other races)</i>								
African American	-0.3183	-6.35	--	--	--	--	-0.1993	-4.22
<i>Housing Tenure (Base category is rental home)</i>								
Own	0.3690	9.48	0.6321	11.08	0.7265	11.71	1.0050	17.81
<i>Household Income (Base category is all lower income levels)</i>								
Greater than \$75K	0.7863	12.91	0.6581	8.89	0.5848	8.29	0.8284	11.80
Number of adults	0.5195	15.57	0.3884	8.27	0.6625	14.02	--	--
Number of full time workers	--	--	0.4139	8.24	--	--	0.4485	9.00
Number of people with more than one job	--	--	-0.3947	-4.12	--	--	--	--
Mean Distance to work (in miles)	--	--	--	--	--	--	0.0115	9.97
<i>Presence of children</i>								
6 to 10 years	-0.2136	-4.48	--	--	--	--	--	--
11 to 15 years	--	--	--	--	0.1985	2.07	--	--
Presence of senior adults	--	--	--	--	-0.5534	-7.87	-0.5617	-8.18
Presence of individual with prolonged medical condition (> 5 yrs)	--	--	-0.3844	-9.21	--	--	--	--
<i>Immigration Status (Base is non-immigrant household)</i>								
Immigrant household	-0.3733	-6.39	-0.3733	-6.39	-0.1957	-3.62	-0.1957	-3.62
Existing vehicle fleet characteristics								
Number of old cars	-1.1195	-12.84	-0.9968	-18.36	-1.0363	-14.35	-0.8437	-12.52
Number of new cars	-2.0845	-32.36	-1.2579	-16.29	-2.3589	-9.24	-0.7159	-10.11
Number of old non-cars	-0.9627	-9.96	-0.7004	-11.51	-1.3573	-15.60	-0.4919	-10.52
Number of new non-cars	-1.5708	-34.02	-1.1944	-8.93	-1.9237	-30.98	-1.7878	-22.35
Accessibility Measures								
Primary arterial roads roadway length within 10 min. (in miles/10 ⁴)	--	--	--	--	-1.7816	-12.15	--	--
Minor arterial roads roadway length within 10 min. (in miles/10 ⁴)	--	--	--	--	-0.9369	-6.03	--	--
Collector roads roadway length within 10 min. (in miles/10 ⁴)	--	--	--	--	0.5306	16.06	--	--
Total manufacturing employment that can be reached within 10 min. (/10 ⁴)	--	--	-2.4896	-2.20	--	--	--	--
Total amount of arts employment that can be reached within 10 min. (/10 ⁴)	-9.2334	-9.33	--	--	--	--	--	--
<i>Spatial Interaction parameter (δ)</i>	0.1872 (3.80)							

TABLE 2 Aggregate-Level Elasticity Effects of the Spatial and Non-Spatial Models

Variables	No Vehicle		Old Car		New Car		Old Non-car		New Non-car	
	Non-spatial	Spatial	Non-spatial	Spatial	Non-spatial	Spatial	Non-spatial	Spatial	Non-spatial	Spatial
Demographics										
<i>Highest Education Attainment in Household (Base is College degree or less)</i>										
Bachelor Degree	3.43	2.90	5.20	4.05	8.53	7.34	5.58	4.97	-83.27	-69.29
Post Graduate	2.71	2.41	4.06	3.50	6.97	5.98	4.61	4.08	-66.52	-57.63
<i>Hispanic Status (Base category is non-Hispanic)</i>										
Hispanic Origin	-5.85	-14.51	-17.21	-39.66	-15.65	-36.84	84.69	197.79	-14.18	-31.90
<i>Race (Base category is all other races)</i>										
African American	8.44	11.67	-45.50	-63.80	22.52	36.01	22.83	35.76	-25.34	-50.94
<i>Housing Tenure (Base category is rental home)</i>										
Own	-29.85	-35.61	1.94	7.49	45.89	60.13	50.88	59.74	101.92	91.74
<i>Household Income (Base category is all lower income levels)</i>										
Greater than \$75K	-39.03	-43.62	56.80	70.00	45.09	40.81	11.23	11.81	88.77	98.06
Number of adults	-23.62	-68.07	34.20	80.83	15.90	20.32	69.54	234.10	-45.62	-88.07
Number of full time workers	-11.24	-45.83	-19.58	-65.85	66.15	240.55	-20.07	-66.73	79.67	312.38
Number of people with more than one job	4.49	7.65	10.02	18.13	-52.21	-92.91	9.83	17.08	13.23	24.81
Mean Distance to work (in miles)	-0.26	-0.59	-0.31	-0.78	-0.71	-1.46	-0.38	-0.77	6.13	13.38
<i>Presence of children</i>										
6 to 10 years	3.82	6.99	-26.48	-50.18	9.96	20.24	11.09	20.64	7.77	14.70
11 to 15 years	-1.96	-5.62	-5.77	-16.28	-5.36	-13.92	28.55	77.49	-4.87	-11.79
Presence of senior adults	9.07	8.62	22.45	21.88	23.76	25.47	-72.84	-68.95	-78.65	-79.93
Presence of individual with prolonged medical condition (> 5 yrs)	4.55	7.07	10.14	16.62	-52.70	-85.69	9.88	15.62	13.23	23.29
<i>Immigration Status (Base is non-immigrant household)</i>										
Immigrant household	16.82	17.65	-28.88	-31.14	-33.84	-36.13	2.16	3.25	-5.46	-4.04
Existing vehicle fleet characteristics										
Number of old cars	45.81	81.34	-60.05	-98.99	-55.64	-97.71	-50.85	-97.31	-46.11	-93.50
Number of new cars	60.12	81.83	-91.74	-100.00	-57.61	-99.71	-93.96	-100.00	4.57	-85.41
Number of old non-cars	39.99	75.88	-50.98	-97.64	-32.43	-86.81	-80.39	-99.93	-8.86	-60.21
Number of new non-cars	61.30	83.14	-75.85	-99.98	-52.10	-99.26	-86.40	-100.00	-89.53	-99.97
Accessibility Measures										
Primary arterial roads roadway length within 10 min. (in miles)	1.81	4.34	4.89	13.70	4.56	12.41	-25.15	-64.82	4.38	11.60
Minor arterial roads roadway length within 10 min. (in miles)	1.18	2.90	3.26	8.89	2.90	7.82	-16.35	-42.05	2.70	7.15
Collector roads roadway length within 10 min. (in miles)	-4.51	-16.76	-10.89	-38.14	-10.52	-34.42	58.78	203.70	-9.49	-31.05
Total amount of manufacturing employment that can be reached within 10 min.	0.23	0.55	0.45	1.08	-2.44	-6.03	0.42	1.02	0.57	1.60
Total amount of arts employment that can be reached within 10 min.	0.65	1.84	-4.41	-12.69	1.83	4.95	1.51	4.92	1.44	3.91