A JOINT TOUR-BASED MODEL OF TOUR COMPLEXITY, PASSENGER ACCOMPANIMENT, VEHICLE TYPE CHOICE, AND TOUR LENGTH

Rajesh Paleti
The University of Texas at Austin
Department of Civil, Architectural & Environmental Engineering
1 University Station C1761, Austin TX 78712-0278
Phone: 512-471-4535, Fax: 512-475-8744
Email: rajeshp@mail.utexas.edu

Ram M. Pendyala
School of Sustainable Engineering and the Built Environment
Arizona State University
Room ECG252, Tempe, AZ 85287-5306
Tel: (480) 965-3549; Fax: (480) 965-0557
Email: ram.pendyala@asu.edu

Chandra R. Bhat
The University of Texas at Austin
Department of Civil, Architectural & Environmental Engineering
1 University Station C1761, Austin TX 78712-0278
Phone: 512-471-4535, Fax: 512-475-8744
Email: bhat@mail.utexas.edu

Karthik C. Konduri
School of Sustainable Engineering and the Built Environment
Arizona State University
Room ECG252, Tempe, AZ 85287-5306
Tel: (480) 965-3549; Fax: (480) 965-0557
Email: karthik.konduri@asu.edu

August 1, 2011

ABSTRACT
Tour-based model systems are increasingly being deployed to microsimulate daily activity-travel patterns of individuals. There are a host of tour attributes of interest that are modeled within these systems. However, a dimension that is often missed is that of vehicle type choice, a variable of considerable importance in the energy consumption and emissions estimation arena. Another issue that arises is that most tour attributes are modeled independently or sequentially with loose coupling across the models, thus ignoring important endogeneity effects that may exist across multiple tour dimensions. This paper considers four key dimensions of tours – tour complexity, passenger accompaniment, vehicle type choice, and tour length – with a view to developing a joint simultaneous equations model system of tour choices while accounting for the presence of correlated unobserved attributes affecting multiple dimensions through appropriate error covariance structures. The paper makes an important methodological contribution by describing and formulating a multi-dimensional joint choice model system capable of accommodating a variety of endogenous variable types (discrete and continuous). The paper makes an important empirical contribution by providing evidence on the nature of the relationships among these tour dimensions of interest within the context of a joint model. The model system is estimated on a sample of tours from the 2009 National Household Travel Survey of the United States. In general, it is found that there is significant evidence of correlated unobserved factors across these tour dimensions and that vehicle type choice affects tour length, a finding that could have important policy implications.

Keywords: simultaneous equations model, activity-travel behavior, multi-dimensional choice model, tour attributes, unobserved attributes, vehicle type choice, tour length, tour complexity, passenger accompaniment

INTRODUCTION

There are a variety of tour attributes of interest in the context of designing and implementing activity- or tour-based microsimulation models of travel. Tour-based model systems generally involve the modeling of all or a subset of tour type, number of intermediate stops, time of day choice, mode choice, intermediate stop purpose, number of individuals on the tour, destination choice for primary and secondary stops, and activity episode duration (Bhat et al, 2004; Vovsha and Bradley, 2006; Bowman and Ben-Akiva, 2001). Model components pertaining to the various choices and dimensions of interest are often linked together to form a sequential chain of models, with potential feedback involving logsums for choice variables where nested logit model forms are used (Wen and Koppelman, 1999).

While the above structures are certainly convenient from a model deployment and application standpoint, they are limited in their ability to simultaneously model the complex inter-relationships among the multiple tour attributes while accounting for the possible presence of correlated unobserved attributes across choice dimensions. The development of simultaneous equations models of activity-travel behavior has been of much interest in the travel behavior research domain for decades for precisely this reason (e.g., Mannering and Hensher, 1987; Kitamura et al, 1996; Pendyala and Bhat, 2004). Simultaneous equations modeling has been motivated by the desire to appropriately represent endogeneity in choice processes where correlated error structures may exist, and thus make travel behavior models more accurately capture behavioral processes at play. Ignoring endogeneity that may exist across choice dimensions that are inter-related with one another results in coefficient estimates that are inconsistent and biased (Mokhtarian and Cao, 2008), with inevitable adverse impacts on the quality of the forecasts provided by such models.

From a methodological perspective, the profession has been limited by the complexity associated with formulating and estimating simultaneous equations models that capture a multitude of choice dimensions in a joint model system. Most simultaneous equations models have been limited to bivariate model systems (e.g., Hamed and Mannering, 1993; Bhat, 1998; Yamamoto and Kitamura, 1999; Golob, 2000; Ye and Pendyala, 2007), either involving two discrete choice variables or a combination of discrete and continuous choice variables. While these models have undoubtedly provided key insights into endogeneity of choice processes, the inability to model more than two choice dimensions simultaneously has made it difficult to account for endogeneity across a multitude of choices that may be made as a package or bundle (Chung and Rao, 2003). The complexity associated with estimating larger multidimensional choice models systems with a mixture of endogenous variable types using classical econometric formulations has led to a stream of literature utilizing structural equations methods (Golob, 2003; Bagley and Mokhtarian, 2002). In structural equations models, multiple endogenous variables may be modeled simultaneously while accounting for the possible presence of significant error covariances. These models have been able to shed considerable light on the complex interactions across multiple activity-travel variables; however, the key issue associated with structural equations models is that they cannot accommodate multinomial choice variables – which happen to be one of the most important variable types in travel modeling (for example, destination choice, mode choice, time of day choice, and activity type choice).

More recently, progress has been made in the multidimensional modeling of choice processes in the activity-travel arena (Pinjari et al, 2011; Eluru et al, 2010). These efforts exploit some of the dramatic advances in choice model specifications and estimation methods that have occurred in the recent past (Bhat and Eluru, 2010; Bhat et al, 2008). These advances make it possible to formulate model specifications that account for complex observed and unobserved inter-relationships that exist among multiple dependent variables and to estimate such model systems without having to resort to simulation-based approaches that quickly become computationally burdensome and potentially imprecise as the dimensionality of the problem increases (Bhat, 2011). The development of methodologies that allow the specification and estimation of complex multi-dimensional choice model systems in simultaneous equations frameworks may be viewed as a major advance with the potential to lead to dramatic breakthroughs in the way activity-based travel model systems are structured and implemented.

This paper aims to further advance the development and estimation of multidimensional choice model systems of activity-travel behavior by considering a bundle of endogeneous variables that characterize tours in activity-based travel model systems. The four attributes considered in this paper are tour complexity, passenger accompaniment, vehicle type choice, and total tour length. While this set of dimensions is certainly not exhaustive by any means, it does represent an important group of choices from a transportation modeling and planning perspective that are likely to be inter-related to one another. Within the context of the emerging energy sustainability and greenhouse gas emissions reduction debates, the modeling of vehicle type choice and tour length is of particular interest as these choices direct impact energy and environmental outcomes. Despite the importance of vehicle type choice in this arena, rarely has vehicle type choice been explicitly modeled in tour-based models. Modeling and tracking vehicle type choice within the larger context of activity-based models can greatly inform emissions inventory models that are able to take advantage of detailed information of vehicle trajectories by type of vehicle. Moreover, from a policy perspective, one can examine the potential (sometimes, unintended) consequences of actions. For example, suppose rebates are instituted for the purchase of fuel efficient vehicles to enhance their presence in the fleet. Individuals can, however, travel farther distances using more fuel efficient vehicles at the same cost as they would travel shorter distances with gas guzzling vehicles. Then, the longer travel distances induced by the acquisition of fuel efficient vehicles (spurred by the policy actions) would, at least in part, negate the benefits associated with encouraging fuel efficient vehicle acquisition in the population. In addition, total vehicle miles of travel could increase, leading to greater levels of congestion and delay. It is these types of complex inter-relationships that can be captured through the estimation and deployment of multi-dimensional choice model systems.

The remainder of this paper is organized as follows. The next section presents a brief overview of the multi-dimensional relationships captured in the model system developed in this paper. The third section presents the modeling methodology in detail. The fourth section offers a description of the data while the fifth section presents model estimation results. The sixth and final section offers concluding thoughts.

MULTI-DIMENSIONAL MODELING OF TOUR ATTRIBUTES

As mentioned earlier, four choice dimensions are considered this paper. They are tour complexity, passenger accompaniment, vehicle type choice, and total tour length. For the model development exercise of this paper, tour complexity is represented by the number of stops made on the tour. A tour itself is defined as a closed chain, with the beginning and ending of the tour being the home location. Only home-based tours are considered in this paper because of the desire to model vehicle type choice and it is presumed that one would have a choice among vehicle types (in a multiple vehicle household) when a tour begins at home. Stop frequency could be represented as an ordered response variable (Bhat and Srinivasan, 2005); however, within the context of this paper, stop frequency is represented as a binary choice variable between the choice of making a one-stop tour or a multiple stop tour. The former may be considered “simple” tours while the latter may be considered “complex” tours. This simplification was done because of the generally low frequency of multiple stop tours in travel survey data sets.

Passenger accompaniment is a variable of much interest because it captures multiple behavioral processes at play. Passenger accompaniment is representative of joint or solo activity engagement, and thus captures (at least in part) interactions among households members. There is increasing recognition of the importance of intra-household interactions in modeling daily activity-travel patterns due to the inevitable linkages and dependencies that exist (Zhang et al, 2005). Children, for example, are often dependent on parents for meeting travel needs (Paleti et al, 2011). Household members often undertake activities jointly, particularly in the context of maintenance and discretionary activities, and this jointness in activity engagement may have important implications for destination choice (tour length), vehicle type choice, and time of day choice. In this paper, passenger accompaniment is represented as a trinary choice variable with possible choice options being a pure solo tour, a pure joint tour (with multiple vehicle occupants throughout the tour), and a partly joint tour (with a single occupant for a part of the tour, and multiple occupants for the other part of the tour).

Vehicle type choice is a variable of much importance and considerable interest from an energy consumption and environmental assessment perspective. However, there is a paucity of research that explicitly addresses vehicle type choice in the context of tour-based model systems. There is a vast body of literature devoted to modeling vehicle ownership. While early research focused heavily on modeling the count of vehicles (Mannering and Winston, 1985), more recent work has provided frameworks for modeling vehicle fleet composition of households with vehicle types defined by body type, make and model, fuel type, and vintage (Bhat and Sen, 2006). In addition, there have been numerous studies that have attempted to model vehicle holding durations, and the timing and nature of vehicle transactions including acquisition, disposal, and replacement of vehicles (Mohammadian and Miller, 2003; Yamamoto et al, 1999). Thus, while there is a base of research that offers methods to model and forecast vehicle ownership by type, there is virtually no research that subsequently uses that information the activity-travel microsimulation process. Vehicle allocation to drivers, and the choice of vehicle for individual tours, are not dimensions that are explicitly simulated, thus limiting the ability to exploit the detailed information output from activity based microsimulation models in estimating energy consumption and emissions inventories. For this reason, the current paper includes vehicle type choice as one of the dimensions in the system. In this paper, for simplicity, vehicle type is represented as a multinomial choice variable with the universe of options being car (auto), sports utility vehicle (SUV), van/minivan, and pick-up truck.

The final choice dimension that is captured in the model system of this paper is destination choice. There is a rich body of evidence on destination choice behavior and spatial processes at play in how people perceive spatial opportunities and choose destinations, while considering the time-space and institutional constraints that govern such choices (Pendyala et al, 2002; Bhat and Zhao, 2002). However, destination choice is a dimension that applies to the individual trip or stop level, and not the tour level, because tours may have multiple destinations associated with multiple stops. In this paper, total tour length is used to capture distances traveled in reaching the one or more destinations on a tour. Total tour length is a continuous variable and is of much interest because it is representative of vehicle miles of travel (VMT), a travel model outcome that is used to quantify total travel and assess impacts on energy and emissions estimates (Shiftan and Suhrbier, 2002).

Thus, the model system in this paper combines a binary choice variable with two multinomial choice variables and one continuous variable. One of the two multinomial choice variables (vehicle type choice) has varying choice sets across choice-makers, depending on the vehicle fleet of the household to which the traveler belongs. The modeling of such a mixture of dependent variable types in a single integrated model system is quite complex, and this paper presents a state-of-the-art methodological framework for doing so by exploiting some of the recent developments in choice modeling and estimation methods. The methodology involves the deployment of new estimation techniques that reduce the dimensionality of the problem, thus eliminating some of the concerns associated with computational burden and imprecision that might arise when adopting simulation-based estimation approaches in the context of large multi-dimensional problems (Bhat, 2011).

Although this section does not constitute a comprehensive review of the literature, it does illustrate the level of interest in the choice dimensions considered in this paper and the need for advances in multi-dimensional integrated choice modeling that would allow the profession to recognize the package or bundle nature of multiple choice processes. While the model system may appear to be a theoretical effort at exercising econometric complexity, the model specification, formulation, and estimation approach presented in this paper offers the potential for dramatic breakthroughs in activity-based travel demand modeling.

MODELING METHODOLOGY

This section presents a detailed description of the modeling methodology developed for estimating a multi-dimensional choice model system involving a mixture of dependent variable types. The model formulation accounts for correlated unobserved factors affecting multiple choice dimensions, and allows the estimation of all model parameters in a single step akin to classic full-information maximum likelihood approaches thus ensuring the use of all information in parameter estimation leading to gains in efficiency. The remainder of this section presents the formulation.

Model Framework

Let there be G nominal variables for an individual, and let g be the index for the nominal variables (g =1, 2, 3,….,G)[1]. In the empirical context of the current paper, G=3 (the nominal variables are accompaniment type, tour type or complexity, and vehicle type). 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). Note that I_g may vary across individuals, but index for individuals is suppressed at this time for ease of presentation. Also, it is possible that some nominal variables do not apply for some individuals, in which case G itself is a function of the individual q. However, the model is developed at the individual level, and so this notational nuance does not appear in the presentation here.

Consider the g^{th} nominal variable and assume that the individual under consideration chooses the alternative m_g. Also, assume the usual random utility structure for each alternative i_g.

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

[1] A nominal variable can be an unordered multinomial response variable or a binary response variable.

where $\boldsymbol{x}_{gi_g}$ is a (K_gx1)-column vector of exogenous attributes, $\boldsymbol{\beta}_g$ 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 $\varepsilon_g = (\varepsilon_{g1}, \varepsilon_{g2}, \ldots\ldots \varepsilon_{gI_g})'$ be $\boldsymbol{\Omega}_g$. As usual, appropriate scale and level normalization must be imposed on $\boldsymbol{\Omega}_g$ for identification (more on this later). Under the utility maximization paradigm, $U_{gi_g} - U_{gm_g}$ must be less than zero for all $i_g \neq m_g$, since the individual chose alternative m_g. Let $y^*_{gi_g m_g} = U_{gi_g} - U_{gm_g} \; (i_g \neq m_g)$, and stack the latent utility differentials into a vector $\boldsymbol{y}^*_{\boldsymbol{g}} = \left[\left(y^*_{g1m_g}, y^*_{g2m_g}, \ldots, y^*_{gI_g m_g} \right)'; i_g \neq m_g \right]$. $\boldsymbol{y}^*_{\boldsymbol{g}}$ has a mean vector of $\boldsymbol{B_g} \, (\boldsymbol{\beta}'_{\boldsymbol{1}} \boldsymbol{z}_{\boldsymbol{g1m_g}}, \boldsymbol{\beta}'_{\boldsymbol{1}} \boldsymbol{z}_{\boldsymbol{g2m_g}}, \ldots. \boldsymbol{\beta}'_{\boldsymbol{1}} \boldsymbol{z}_{\boldsymbol{gI_g m_g}})'$, where $\boldsymbol{z}_{\boldsymbol{gi_g m_g}} = \boldsymbol{x}_{\boldsymbol{gi_g}} - \boldsymbol{x}_{\boldsymbol{gm_g}}, i_g = 1,2,\ldots I_g; i_g \neq m_g$. To obtain the covariance matrix of $\boldsymbol{y}^*_{\boldsymbol{g}}$, define $\boldsymbol{M}_{\mathbf{g}}$ as an (I_g-1) x 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. Then, one may write:

$$\boldsymbol{y}^*_{\boldsymbol{g}} \sim N\left(\boldsymbol{B}_{\boldsymbol{g}}, \boldsymbol{\Sigma}^*_{\mathbf{g}}\right), \text{where } \boldsymbol{\Sigma}^*_{\mathbf{g}} = \boldsymbol{M}_{\boldsymbol{g}} \boldsymbol{\Omega}_{\mathbf{g}} \boldsymbol{M}'_{\boldsymbol{g}}. \tag{2}$$

The discussion above focuses on a single nominal variable g. When there are G nominal variables, consider the stacked $\left[\sum_{g=1}^{G} (I_g - 1)\right] \times 1-$vector $\boldsymbol{y}^* = \left[\left(\boldsymbol{y}^{*'}_{\boldsymbol{1}}, \boldsymbol{y}^{*'}_{\boldsymbol{2}}, \ldots\ldots \boldsymbol{y}^{*'}_{\boldsymbol{G}} \right)' \right]$, each of whose element vectors is formed by differencing utilities of alternatives from the chosen alternative m_g for the g^{th} nominal variable. Next, one may write:

$\boldsymbol{y}^* \sim N(\boldsymbol{B}, \boldsymbol{\Sigma}^*)$, where $\boldsymbol{B} = \left(\boldsymbol{B}'_{\boldsymbol{1}}, \boldsymbol{B}'_{\boldsymbol{2}}, \ldots \boldsymbol{B}'_{\boldsymbol{G}} \right)'$ and $\boldsymbol{\Sigma}^*$ is a $\left[\sum_{g=1}^{G} (I_g - 1)\right] * \left[\sum_{g=1}^{G} (I_g - 1)\right]$ matrix as follows:

$$\boldsymbol{\Sigma}^* = \begin{bmatrix} \boldsymbol{\Sigma}^*_{\mathbf{1}} & \boldsymbol{\Sigma}^*_{\mathbf{12}} & . & . & . & \boldsymbol{\Sigma}^*_{\mathbf{1G}} \\ \boldsymbol{\Sigma}^*_{\mathbf{21}} & \boldsymbol{\Sigma}^*_{\mathbf{2}} & . & . & . & \boldsymbol{\Sigma}^*_{\mathbf{2G}} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \boldsymbol{\Sigma}^*_{\mathbf{G1}} & \boldsymbol{\Sigma}^*_{\mathbf{G2}} & . & . & . & \boldsymbol{\Sigma}^*_{\mathbf{G}} \end{bmatrix} \tag{3}$$

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

Now, assume that, in addition to the G nominal variables, there are H continuous variables $(y_1, y_2, \ldots\ldots y_H)$ with an associated index h (h = 1, 2,…H). Let $y_h = \boldsymbol{\gamma}'_{\boldsymbol{h}} \boldsymbol{s}_h + \eta_h$ in the usual linear regression fashion. Stacking the H continuous variables into a (H x 1)-vector $\boldsymbol{y}$, one may write $y = N(\boldsymbol{c}, \boldsymbol{\Sigma})$, where $\boldsymbol{c} = \left(\boldsymbol{\gamma}'_{\boldsymbol{1}} \boldsymbol{s}_{\boldsymbol{1}}, \boldsymbol{\gamma}'_{\boldsymbol{2}} \boldsymbol{s}_{\boldsymbol{2}}, \ldots\ldots \boldsymbol{\gamma}'_{\boldsymbol{H}} \boldsymbol{s}_{\boldsymbol{H}} \right)'$, and $\boldsymbol{\Sigma}$ is the covariance matrix of $\boldsymbol{\eta} = \left(\eta_1, \eta_2, \ldots\ldots \eta_H \right)'$. The variance of $\tilde{\boldsymbol{y}} = \left(\boldsymbol{y}^{*'}, \boldsymbol{y}' \right)'$ can be written as:

$$\operatorname{Var}(\tilde{\boldsymbol{y}}) = \tilde{\boldsymbol{\Lambda}} = \begin{bmatrix} \boldsymbol{\Sigma}^* & \boldsymbol{\Sigma}_{\mathbf{y}^*\mathbf{y}} \\ \boldsymbol{\Sigma}'_{\mathbf{y}^*\mathbf{y}} & \boldsymbol{\Sigma} \end{bmatrix}, \tag{4}$$

where $\boldsymbol{\Sigma}_{\mathbf{y}^*\mathbf{y}}$ is a $\left[\sum_{g=1}^{G}(I_g-1)\right]xH$ matrix capturing covariance effects between the $\boldsymbol{y}^*$ vector and the $\boldsymbol{y}$ vector. The conditional distribution of $\boldsymbol{y}^*$, given $\boldsymbol{y}$, is multivariate normal with mean $\tilde{\boldsymbol{B}} = \boldsymbol{B} + \boldsymbol{\Sigma}_{\mathbf{y}^*\mathbf{y}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{c})$ and variance $\tilde{\boldsymbol{\Sigma}}^* = \boldsymbol{\Sigma}^* - \boldsymbol{\Sigma}_{\mathbf{y}^*\mathbf{y}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}'_{\mathbf{y}^*\mathbf{y}}$. The basis for the construction of the $\tilde{\boldsymbol{\Lambda}}$ matrix will be different for different individuals, since the chosen alternative for each nominal variable will, in general, be different across individuals. At the same time, it must be ensured that $\tilde{\boldsymbol{\Lambda}}$ across individuals is derived from a common covariance matrix $\boldsymbol{\Lambda}$ for the original $\left[\left(\sum_{g=1}^{G} I_g\right)+H\right]$-error term vector $(\boldsymbol{\varepsilon}',\eta')'$, subject to identification considerations $[\varepsilon = (\varepsilon_1', \varepsilon_2', \ldots \varepsilon_G')']$. Also, the overall matrix $\tilde{\boldsymbol{\Lambda}}$ needs to be positive definite (as will be discussed later).[2]

Next, let $\boldsymbol{\theta}$ be the collection of parameters to be estimated: $\boldsymbol{\theta} = [\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \ldots \boldsymbol{\beta_G}; Vech(\boldsymbol{\Sigma}^*); \boldsymbol{\gamma_1}, \boldsymbol{\gamma_2}, \ldots \boldsymbol{\gamma_H}; Vech(\boldsymbol{\Sigma}); Vech(\boldsymbol{\Sigma}_{\mathbf{y}^*\mathbf{y}})]$, where Vech($\boldsymbol{\Sigma}$) represents the vector of upper triangle elements of $\boldsymbol{\Sigma}$. Then the likelihood function for the individual may be written as:

$$L(\boldsymbol{\theta}) = \phi_H(\boldsymbol{y}-\boldsymbol{c}\,|\,\boldsymbol{\Sigma}) \times F_{\tilde{G}}\left(\tilde{\boldsymbol{B}}, \tilde{\boldsymbol{\Sigma}}^*\right), \tag{5}$$

where $\phi_H(.\,|\,\boldsymbol{\Sigma})$ is the H-dimensional normal density with mean 0 and covariance matrix $\boldsymbol{\Sigma}$, and $F_{\tilde{G}}(.,.)$ is the $\tilde{G} = \left(\left[\sum_{g=1}^{G}(I_g-1)\right]\right)$-dimensional normal cumulative distribution function.

The above likelihood function involves the evaluation of a $\tilde{G}$-dimensional integral for each individual, which can be very expensive if there are several nominal variables or if each nominal variable can take a large number of values or a combination of the two. So, the Maximum Approximated Composite Marginal Likelihood (MACML) approach of Bhat (2011), in which the likelihood function only involves the computation of univariate and bivariate cumulative distributive functions, is used in this paper.

The MACML Estimation Approach

Consider the following (pairwise) composite marginal likelihood function formed by taking the products (across the G nominal variables) of the joint pairwise probability of the chosen alternatives m_g and m_l for the g^{th} and l^{th} nominal variables for an individual.

$$L_{CML}(\boldsymbol{\theta}) = \phi_H(\boldsymbol{y}-\boldsymbol{c}\,|\,\boldsymbol{\Sigma}) * \prod_{g=1}^{G-1}\prod_{l=g+1}^{G}\Pr(d_{i_g}=m_g, d_{i_l}=m_l), \tag{6}$$

where d_{i_g} is an index for the individual's choice for the g^{th} nominal variable, and m_g is the actual chosen alternative for the g^{th} nominal variable. One can write:

[2] Note that if $\tilde{\boldsymbol{\Lambda}}$ is positive definite, then it immediately implies that $\boldsymbol{\Sigma}^*$ (and each of $\boldsymbol{\Sigma}_1^*, \boldsymbol{\Sigma}_2^*, \ldots, \boldsymbol{\Sigma}_G^*$) as well as $\boldsymbol{\Sigma}$ are all positive definite because of the property that any principal square sub-matrix of a positive definite matrix is also positive definite.

$$\Pr(d_{i_g} = m_g, d_{i_l} = m_l) = F_{\tilde{G}_{gl}}(\mathbf{\Delta_{gl}}\tilde{\boldsymbol{B}}, \mathbf{\Delta_{gl}}\tilde{\mathbf{\Sigma}}^{*}\mathbf{\Delta}'_{\mathbf{gl}}), \tag{7}$$

where $\tilde{G}_{gl} = I_g + I_l - 2$ (I_g is the number of alternatives for the g^{th} nominal variable) and $\mathbf{\Delta_{gl}}$ is a $\tilde{G}_{gl} * \tilde{G}$ -selection matrix with an identity matrix of size ($I_g - 1$) occupying the first ($I_g - 1$) rows and the $\left[\sum_{j=1}^{g-1}(I_j - 1) + 1\right]^{th}$ through $\left[\sum_{j=1}^{g}(I_j - 1)\right]^{th}$ columns (with the convention that $\sum_{j=1}^{0}(I_j - 1) = 0$, and another identity matrix of size ($I_l - 1$) occupying the last ($I_l - 1$) rows and the $\left[\sum_{j=1}^{l-1}(I_j - 1) + 1\right]^{th}$ through $\left[\sum_{j=1}^{l}(I_j - 1)\right]^{th}$ columns. The net result is that the pairwise likelihood function now only needs the evaluation of a $\tilde{G}_{gl}$-dimensional cumulative normal distribution function (rather than the $\tilde{G}$ -dimensional cumulative distribution function in the maximum likelihood function). This can lead to substantial computation efficiency. However, in cases where there are several alternatives for one or more nominal variables, the dimension $\tilde{G}_{gl}$ can still be quite high. This is where the use of an analytic approximation of the multivariate normal cumulative distribution (MVNCD) function, as shown in Bhat (2011), is convenient. The resulting maximum approximated composite marginal likelihood (MACML) of Bhat (2011), which combines the CML approach with the analytic approximation for the MVNCD function evaluation, is solely based on bivariate and univariate cumulative normal computations. The MACML approach can be applied using a simple optimization approach for likelihood estimation. It also represents a conceptually simpler alternative to simulation techniques. Also, the MACML estimator $\hat{\boldsymbol{\theta}}_{MACML}$ is asymptotically normal distributed with mean $\boldsymbol{\theta}$ and covariance matrix given by the inverse of the Godambe's (1960) sandwich information matrix $G(\boldsymbol{\theta})$ (Zhao and Joe, 2005):

$$V_{MACML}(\boldsymbol{\theta}) = [G(\boldsymbol{\theta})]^{-1} = H(\boldsymbol{\theta})[J(\boldsymbol{\theta})]^{-1}[H(\boldsymbol{\theta})], \tag{8}$$

where $H(\boldsymbol{\theta})$ and $J(\boldsymbol{\theta})$ take the following form:

$$H(\boldsymbol{\theta}) = E\left[-\frac{\partial^2 \log L_{MACML}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right] \text{ and } J(\boldsymbol{\theta}) = E\left[\left(\frac{\partial \log L_{MACML}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\left(\frac{\partial \log L_{MACML}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right)\right]$$

$H(\boldsymbol{\theta})$ and $J(\boldsymbol{\theta})$ can be estimated in a straightforward manner at the MACML estimate $\hat{\boldsymbol{\theta}}_{\boldsymbol{MACML}}$ as follows (introducing q as the index for individuals):

$$\hat{H}(\hat{\boldsymbol{\theta}}) = -\left[\sum_{q=1}^{Q} \frac{\partial^2 \log L_{MACML,q}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]_{\hat{\theta}_{CML}}, \text{ and}$$

$$\hat{J}(\hat{\boldsymbol{\theta}}) = \sum_{q=1}^{Q}\left[\left(\frac{\partial \log L_{MACML,q}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\left(\frac{\partial \log L_{MACML,q}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right)\right]_{\hat{\theta}_{CML}}. \tag{9}$$

Ensuring Identification and Positive Definiteness

There are two important issues that need to be dealt with during estimation, each of which is discussed in this section.

Identification

The estimated model needs to be theoretically identified. As discussed earlier, in a model with a nominal dependent variable, only utility differences matter. Suppose one considers utility differences with respect to the first alternative for each of the G nominal variables. Then, the analyst can restrict the variance term of the top left diagonal of the resulting covariance matrix (say $\breve{\mathbf{\Sigma}}_g^*$) of utility differences to 1 to account for scale invariance. However, note that the matrix $\breve{\mathbf{\Sigma}}_g^*$ is different from the matrix $\mathbf{\Sigma}_g^*$ which corresponds to the covariance of utility differences taken with respect to the chosen alternative for the individual. Next, create a matrix of dimension $\left[\sum_{g=1}^{G}(I_g-1)\right]\times\left[\sum_{g=1}^{G}(I_g-1)\right]$ similar to that of $\mathbf{\Sigma}^*$ in Equation (3), except that the matrix is expressed in terms of utility differences with respect to the first alternative for each nominal variable:

$$\breve{\mathbf{\Sigma}}^* = \begin{bmatrix} \breve{\mathbf{\Sigma}}_1^* & \breve{\mathbf{\Sigma}}_{12}^* & . & . & . & \breve{\mathbf{\Sigma}}_{1G}^* \\ \breve{\mathbf{\Sigma}}_{21}^* & \breve{\mathbf{\Sigma}}_2^* & . & . & . & \breve{\mathbf{\Sigma}}_{2G}^* \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \breve{\mathbf{\Sigma}}_{G1}^* & \breve{\mathbf{\Sigma}}_{G2}^* & . & . & . & \breve{\mathbf{\Sigma}}_G^* \end{bmatrix} \tag{10}$$

Further, construct an enhanced covariance matrix that includes the covariance matrix $\mathbf{\Sigma}$ of $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots\ldots\eta_H)$ as follows:

$$\breve{\mathbf{\Omega}} = \begin{bmatrix} \breve{\mathbf{\Sigma}}^* & \breve{\mathbf{\Psi}} \\ \breve{\mathbf{\Psi}}' & \mathbf{\Sigma} \end{bmatrix}, \tag{11}$$

where the $\left[\sum_{g=1}^{G}(I_g-1)\right]\times H-$matrix $\breve{\mathbf{\Psi}}$ contains the covariances between the latent utility differentials (taken with respect to the first alternative) and the $\boldsymbol{y}$ vector. All elements of the matrix $\breve{\mathbf{\Omega}}$ are identifiable, and are the ones estimated. In the general case, this allows the estimation of $\sum_{g=1}^{G}\left(\frac{I_g*(I_g-1)}{2}-1\right)$ terms across all the G nominal variables (originating from $\left(\frac{I_g*(I_g-1)}{2}-1\right)$ terms embedded in each $\breve{\mathbf{\Sigma}}_g^*$ matrix; g=1,2,…G), $\sum_{g=1}^{G-1}\sum_{l=g+1}^{G}(I_g-1)\times(I_l-1)$ covariance terms in the off-diagonal matrices of the $\breve{\mathbf{\Sigma}}^*$ matrix characterizing the dependence between the latent utility differentials (with respect to the first alternative) across the nominal variables (originating from $(I_g-1)\times(I_l-1)$ estimable covariance terms within each off-diagonal matrix $\breve{\boldsymbol{\Sigma}}^*_{\mathbf{gl}}$ in $\breve{\boldsymbol{\Sigma}}^*$), $\left[\sum_{g=1}^{G}(I_g-1)\right]\times H$ covariance terms in $\breve{\boldsymbol{\Psi}}$ for the dependence between the latent utility differentials and the linear regression errors, and the $[H\times(H+1)]/2]$ covariance terms in $\boldsymbol{\Sigma}$.

To construct the general covariance matrix $\boldsymbol{\Lambda}$ for the original $\left[\left(\sum_{g=1}^{G} I_g\right)+H\right]$-error term vector $(\boldsymbol{\varepsilon}',\eta')'$, while also ensuring all parameters are identifiable, zero row and column vectors are inserted for the first alternatives of each nominal variable in $\breve{\boldsymbol{\Omega}}$. To do so, define a matrix ***D*** of size $\left[\left(\sum_{g=1}^{G} I_g\right)+H\right]\times\left[\left(\sum_{g=1}^{G}(I_g-1)\right)+H\right]$. The first I_1 rows and (I_1-1) columns correspond to the first nominal variable. Insert an identity matrix of size (I_1-1) after supplementing with a first row of zeros into this first I_1 rows and (I_1-1) columns of ***D***. The rest of the columns for the first I_1 rows and the rest of the rows for the first (I_1-1) columns take a value of zero. Next, rows (I_1+1) through (I_1+I_2) and columns (I_1) through (I_1+I_2-2) correspond to the second nominal variable. Again position an identity matrix of size (I_2-1) after supplementing with a first row of zeros into this position. Continue this for all *G* nominal variables. Finally, insert an identity matrix of size *H* into the last *H* rows and *H* columns of the matrix ***D*** (with all other columns of these last *H* rows and all other rows of these last *H* columns receiving a value of zero). Thus, for the case with two nominal variables, one nominal variable with 3 alternatives and the second with four alternatives, and two continuous variables, the matrix ***D*** takes the form shown below:

$$\left[\begin{array}{cc:ccc:cc}
0 & 0 & 0 & 0 & 0 & 0 & 0\\
1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & 0 & 0\\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 0\\
\hdashline
0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]_{9*7}$$

Then, the general covariance matrix may be developed as $\boldsymbol{\Lambda}=\mathbf{D}\breve{\boldsymbol{\Omega}}\mathbf{D}'$. All parameters in this matrix are identifiable by virtue of the way this matrix is constructed based on utility differences and, at the same time, it provides a consistent means to obtain the covariance matrix $\tilde{\boldsymbol{\Lambda}}$ that is needed for estimation (and is with respect to each individual's chosen alternative for each nominal variable). Specifically, define a matrix ***M*** of size

$\left[\left(\sum_{g=1}^{G}(I_g-1)\right)+H\right]\times\left[\left(\sum_{g=1}^{G}I_g\right)+H\right]$. The first (I_1-1) rows and I_1 columns correspond to the first nominal variable. Insert an identity matrix of size (I_1-1) after supplementing with a column of '-1' values in the column corresponding to the chosen alternative. The rest of the columns for the first (I_1-1) rows and the rest of the rows for the first I_1 columns take a value of zero. Next, rows (I_1) through (I_1+I_2-2) and columns (I_1+1) through (I_1+I_2) correspond to the second nominal variable. Again position an identity matrix of size (I_2-1) after supplementing with a column of '-1' values in the column corresponding to the chosen alternative. Continue this procedure for all *G* nominal variables. Finally, insert an identity matrix of size *H* into the last *H* rows and *H* columns of the matrix ***M***. With the matrix ***M*** as defined, the covariance matrix $\tilde{\mathbf{\Lambda}}$ for any individual is given by $\tilde{\mathbf{\Lambda}}=\mathbf{M\Lambda M'}$.

Positive Definiteness

The matrix $\tilde{\mathbf{\Lambda}}$ for any individual has to be positive definite. The simplest way to guarantee this is to ensure that the matrix $\breve{\mathbf{\Omega}}$ is positive definite (recall that this is the covariance matrix for the utility differentials with respect to the first alternative). To do so, the Cholesky matrix of $\breve{\mathbf{\Omega}}$ may be used as the matrix of parameters to be estimated. However, note that the top diagonal element of each $\breve{\mathbf{\Sigma}}_g^*$ is normalized to one for identification, and this restriction should be recognized when using the Choleski factor of $\breve{\mathbf{\Omega}}$. This can be achieved by appropriately parameterizing the diagonal elements of the Cholesky decomposition matrix. Thus, consider the lower triangular Choleski matrix $\breve{\boldsymbol{L}}$ of the same size as $\breve{\mathbf{\Omega}}$. Whenever a diagonal element (say the *kk*th element) of $\breve{\mathbf{\Omega}}$ is to be normalized to one, the first element in the corresponding row of $\breve{\boldsymbol{L}}$ is written as $\sqrt{1-\sum_{j=2}^{k}d_{kj}^2}$, where the d_{kj} elements are the Cholesky factors that are to be estimated. With this parameterization, $\breve{\mathbf{\Omega}}$ obtained as $\breve{\boldsymbol{L}}\breve{\boldsymbol{L}}'$ is positive definite and adheres to the scaling conditions. Using this, one constructs $\mathbf{\Lambda}$, and subsequently obtains $\tilde{\mathbf{\Lambda}}$ as discussed in the previous section. The resulting $\tilde{\mathbf{\Lambda}}$ is positive definite, since it is constructed to be consistent with $\breve{\mathbf{\Omega}}$, which is positive-definite.

DATA DESCRIPTION

The data for this study is derived from the 2009 National Household Travel Survey of the United States. This survey collects detailed socio-economic, demographic, travel, and vehicle information for a sample of households in the nation. Each trip (involving a personal automobile use) is tagged with the identity of the vehicle in the household that was used for the trip. Trip chains or tours can be easily constructed from the trip file. For this study, all closed loops or chains that began and ended at home were constructed as home-based tours and those that began and ended at work were constructed as work-based tours. As the analysis involves the choice of vehicle type, only tours undertaken by individuals in households that have multiple vehicles were chosen for analysis. Presumably, individuals in households with zero or one vehicle do not have a choice in vehicle usage. In addition, as vehicle type choice is likely to be limited at the home anchor, only home-based tours were selected for inclusion in the analysis sample. As tours involving journey to and from work are often time-space constrained and may involve aspects that constrain vehicle type choice (e.g., service workers who need pick-up truck for transporting tools of the trade), only home-based non-work tours were considered for analysis. Finally, the analysis was limited to home-based non-work vehicle tours undertaken on regular weekdays – Monday through Thursday – by individuals aged 15 years or over. These filtering criteria resulted in a total sample size of 66,030 home-based non-work tours suitable for analysis. For ease of computation, and to avoid the artificial inflation of test statistics that may lead to erroneous inferences, a random sample of 6,478 tours (nearly 10 percent) were selected for model estimation. Table 1 provides descriptive statistics for the subsample of HBNW tours. Each HBNW tour involved an average of 1.7 stops with average travel duration of 37 minutes and average tour length of 15.7 miles. On average, there were about 1.7 persons on each tour, reflecting the higher vehicle occupancies often associated with non-work travel. Each household in the subsample comprised of nearly three persons with one child. Most of the households in the sample (68 percent) reside in urban areas. There is a slightly higher percentage (54 percent) of females than males. This may be an artifact of limiting the analysis to non-work tours (e.g., involving household maintenance, serve-child) which may be undertaken more so by women than men. As the analysis is limited to tours undertaken by individuals in multi-vehicle households, the average vehicle ownership for the analysis sample is quite high at 2.8 cars per household. Nearly 20 percent of households report having four or more cars, reasonably consistent with the fact that the sample has 34 percent of households with four or more persons.

Table 1 also shows the distribution of vehicle types chosen for the tours in the estimation sample. First, the distribution is chosen for all tours. It is found that 42 percent of all tours are undertaken by auto, 25 percent by SUV, 14 percent by van, and 19 percent by pick-up truck. While these percentages might suggest that individuals are more inclined to choose cars and SUVs for travel, that may not necessarily be true because these percentages do not account for the differential availability of different vehicle types in the fleet. When one controls for vehicle availability in the fleet, then it is found that auto, van, and SUV all enjoy virtually identical probabilities of being chosen at about 50 percent. Only pick-up truck has a lower probability of being chosen at about 32 percent. In other words, when auto, van, or SUV is available in the household fleet, each of these vehicles has a one-in-two chance of being chosen for a tour. When pick-up truck is available in the fleet, the probability of it being selected for a tour is only about one-in-three. The percent of all tours undertaken by auto is greater than that for all other vehicle types simply because it is more available (present) in the household fleets. This demonstrates the importance of accounting for differential choice set composition when estimating models of vehicle type choice and drawing inferences regarding vehicle choices. Additional detailed statistics on tour attributes by vehicle type are shown in Table 2. The table is rather self-explanatory with descriptive statistics consistent with expectations. Average vehicle occupancy, for example, is greater for tours undertaken by van and SUV, presumably because these vehicles are likely to be owned and used by larger size households.

MODEL ESTIMATION RESULTS

This section presents a detailed discussion of the model estimation results. A variety of models were estimated to understand the nature of relationships among the four tour attributes considered in this paper.

Structure of Relationships Among Endogenous Variables

Before proceeding to a discussion of the estimates of coefficients and error covariances, it may be beneficial to consider behavioral hypotheses governing the nature of relationships among the endogenous variables. Figure 1 presents a flow chart depicting the structure of relationships that guided the model specification and estimation. Socio-economic and demographic attributes are assumed to affect all endogenous variables. Among the endogenous variables themselves, passenger accompaniment (which is an endogenous variable because it is a function of explanatory variables) is assumed to impact tour complexity. In the context of a joint or partly joint tour, it is more likely that additional stops will be made to serve the needs of the passenger or to engage in a series of joint activities (for example, eat dinner at a restaurant and then go to the movies). In addition, however, passenger accompaniment may also affect vehicle type choice. When there are multiple individuals involved in a trip, then the larger vehicle may be chosen for reasons of comfort. Finally, passenger accompaniment may also affect tour length. When a joint activity is involved, or a passenger needs to be dropped off or picked up, then destinations are often dictated by the collective needs and desires of the multiple occupants. This may result in traveling to and from locations that are farther away than would otherwise be the case. As such, passenger accompaniment is postulated as affecting all three of the other endogenous variables.

Next, consider tour complexity which is a binary choice variable indicating whether the tour involved a single stop or multiple stops. Tour complexity is assumed to impact both vehicle type choice and tour length. In the context of vehicle type choice, it is possible that larger and more comfortable vehicle types will be used for multiple stop tours. Also, multi-stop tours are likely to be of longer distance because of the need to travel to multiple locations. As multi-stop tours are longer in distance, two possibilities arise. Tour length may, in turn, influence vehicle type choice. First, if tour length is longer (because the tour is complex), individuals may choose the more fuel efficient vehicle type to reduce travel costs associated with traveling long distances. On the other hand, if a person would like to increase comfort levels during a long multi-stop tour, then the larger vehicle type may be chosen to undertake the trip.

In other words, the relationship between the last two variables is subject to debate. While the flowchart shows vehicle type choice affecting tour length, it is entirely possible that tour length affects vehicle type choice. If vehicle type choice affects tour length, then one is implying that people make conscious choices regarding destinations (miles of travel) depending on the nature of the vehicle being used. If a person is using a small fuel efficient car, would the person visit farther destinations and travel more miles because it is possible to do so at lower cost than if a gas guzzling vehicle were used? Or would the person visit close-by destinations and reduce mileage because traveling long distances in the small fuel efficient vehicle is uncomfortable? Alternatively, if the traveler has to visit destinations farther away, then would the fuel efficient vehicle be chosen to keep costs down? Or would a large gas guzzling vehicle be used to maximize comfort levels on the tour? In a previous study, Konduri et al (2011) found that a model in which vehicle type choice affects tour length is statistically superior to a model specification in which tour length affects vehicle type choice. While that finding is clear and intuitive, as vehicle type choice (and allocation of vehicles in a household to drivers) is likely to be a longer term decision relative to tour length choices, the study did not account for the possible endogeneity of passenger accompaniment and tour complexity. Both of these dimensions were treated as exogenous variables, potentially resulting in erroneous inferences regarding the direction of the relationship between vehicle type choice and tour length. This study offers the opportunity to further explore the nature of the relationship between these two variables while accounting for the endogeneity of passenger accompaniment and tour complexity.

Model Results

Model estimation results are presented in Table 3. The constants in the model of passenger accompaniment suggest that partly joint tours are the least likely tour type (all other things being equal) and solo tours are the most likely. A host of socio-economic attributes impact tour accompaniment. Tours undertaken by individuals in households with larger household size (relative to number of vehicles) or larger number of children (relative to the number of drivers) are more likely to be joint tours than solo tours as evidenced by the positive coefficients on these variables. In particular, the presence of children appears to induce partly joint tours, a finding that is consistent with the notion that such households undertake serve-child tours where the child accompanies the parent for a part of the tour and the driver is alone for the remainder of the tour. Males are less likely to undertake joint tours, possibly because females are more likely to take care of household responsibilities and chauffeuring of children. Those 18 years and younger are likely to undertake joint tours, but less likely to undertake partly joint tours; this finding is consistent with expectations, considering that the sample is restricted to those of driving age. These individuals probably drive themselves in solo tours as opposed to needing a partly joint tour involving a pick-up/drop-off. They are, however, more likely to engage in full joint tours in consort with other household members. Part-time employment is associated with greater participation in partly joint tour; perhaps part-time employees are more able to undertake serve-passenger and serve-child activities on behalf of the household leading to a greater prevalence of these partly joint tours for this demographic.

In general, complex tours are less likely to occur than simple tours (all other things being equal) as evidenced by the negative constant for the complexity utility equation (although the coefficient is not statistically significant). A rather surprising finding is that joint and partly joint tours are less likely to be complex than solo tours. One would have expected these tour types, that involve multiple passengers, to be more complex. On the other hand, it is possible that this finding is quite intuitive. When multiple passengers are involved, then the driver or any one individual may not be able to undertake a series of activities on a tour that are of no interest or relevance to other passengers on the tour. The individuals on the joint tour are collectively going to a certain location, undertaking a joint activity, and then returning to base. Only the activity that is of interest and relevance to the entire group is visited. Those who work full time are less likely to engage in complex tours, presumably because of time constraints associated with full time employment. Younger individuals are less likely to undertake complex tours. Males are also less likely to undertake complex tours, suggesting that females have more complex activity-travel patterns as they shoulder a greater share of household responsibilities.

The vehicle type choice model is presented next within Table 3. The constants are all very significant, with the auto and SUV having the highest constants suggesting that these two vehicle types tend to get used more often than others when they are in the choice set. Van also has a positive coefficient suggesting that it is used more than the pick-up truck which is the base alternative. As expected, joint tours are more likely to be undertaken by van or the larger vehicle type, suggesting that the more comfortable (larger) vehicle type is chosen when multiple occupants are involved. In the case of a partly joint tour, the auto is less likely to be used among all vehicle types. Males are likely to use the pick-up truck (when it is available in the household fleet) compared to all other vehicle types when they are present in the fleet. This finding is consistent with expectations and illustrative of the strong gender influence in the pick-up truck market. Those aged 65 years or over and those with children in the household are most likely to use van for tours. The older age group may enjoy the comfort and driving ease of a van, and may have more use for the van as they transport grand children or grown children. It is not surprising that the presence of children is associated with a positive impact on van use; households with child transport duties would likely enjoy the space and comfort of van for chauffeuring duties. Households in non-urban areas are more likely to use a car or a pick-up truck in comparison to van and SUV. This is also consistent with expectations in that the van and SUV are probably not the most preferred vehicle types in rural areas. It is interesting to note that tour complexity does not directly enter the equation of vehicle type choice. It appears that tour complexity does not truly directly influence vehicle type choice; rather it is the accompaniment that influences vehicle type choice.

Finally, the model of tour length shows that accompaniment, complexity, and vehicle type affect tour length. In other words, tour length is affected by all other tour attributes. According to the model estimation results presented in the last part of Table 3, joint tours are likely to be of longer length. This is consistent with the notion that tours involving multiple people might be longer in distance in an effort to find destinations that satisfy the desires of all individuals on the tour. Similarly, tour complexity also adds significantly to tour length. As one adds stops to a tour, it is natural to expect tour length to increase as the addition of each stop entails some additional travel distance. Among the vehicle types, vans tend to have the longer tour length, presumably because vans are comfortable for transporting household members or undertaking joint activities. It is somewhat surprising to see that cars are next in line in terms of a positive impact on tour length, while sports utility vehicles and pick-up trucks show the lowest impact on tour length. This finding is a key sign that people are making a conscious trade-off in the distance traveled by different vehicle types. Both sports utility vehicles and pick-up trucks generally have the poorest fuel economy among all vehicle types. The model is indicating that both of these vehicle types are associated with the shortest tour lengths relative to car and van vehicle types (both of which tend to have better fuel economy). It appears that individuals are making conscious decisions involving trade-offs between travel cost and miles of travel. If a large gas guzzling vehicle is used, then the individual may attempt to consciously find locations that are closer in distance to reduce travel costs. Of course, such trade-offs can be exercised only in the context of non-work tours/travel.

The impacts of socio-economic and demographic attributes on tour length are in line with expectations. As household size (relative to number of vehicles) increases, the tour length decreases. This may be reflective of the vehicle availability constraints that the household has to deal with. In households where household size is large relative to number of vehicles, individuals who take the vehicle to undertake a tour may have to return quickly so that another household member can use the same vehicle. This compels travelers to undertake short tours and minimize travel time. The lower number of vehicles relative to household size may also be reflective of a lower income level; individuals in such households may purposefully undertake shorter distance tours to save on travel costs. As the number of children increases, individuals tend to make shorter tours. This is presumably due to two reasons. First, if the children are accompanying the tour maker, then the individual may choose to complete errands quickly by undertaking shorter tours in order to avoid taxing the patience of the children. Second, if the children are not accompanying the individual on the tour, then it might be necessary for the individual to quickly conclude the tour and return home to tend to children. It is also possible that children have schedule constraints that compel the traveler to undertake shorter tours. Males tend to make longer tours suggesting that females make shorter tours visiting destinations more closely located to the home base. Those with higher education undertake longer tours, perhaps because they have higher income levels, or are more aware of desirable destinations for non-work activities. As expected, those in non-urban areas undertake longer tours; this is likely due to the lower levels of accessibility to destinations enjoyed by such households. Lower income individuals make shorter tours as do individuals 65 years of age and over. Older individuals may not be comfortable traveling long distances. Those with flexible work start time, and thus less rigid time-space constraints associated with work, are found to engage in short tour lengths. This is presumably because these individuals do not have to chain multiple activities into longer multi-stop tours in the quest for efficiency; instead, they can engage in a larger number of short tours. Indeed, the work time flexibility is negatively associated with complex tour formation.

Model Assessment

This section presents a brief assessment of the joint model estimated and presented in this paper. The log likelihood of the final joint model accounting for all potential correlations is significantly better than that of the independent model where all dimensions are estimated separately. The log-likelihood value for the joint model is -23487.87 while that for the independent model ignoring error correlations is -23535.74. The likelihood ratio test statistic is found to be 95.75 with 12 degrees of freedom. This value is considerably greater than the critical χ^2 value of 21.03 at 12 degrees of freedom, suggesting that the joint model offers a statistically superior fit at a 0.05 level of significance. This finding of improved goodness-of-fit of the joint model is the first indication that there may significant error correlations that contribute to a poorer fit in the independent model where they are ignored.

The covariance matrix $\breve{\boldsymbol{\Omega}}$ for the utility differentials[3] with respect to the first alternative corresponding to the Cholesky matrix $\breve{\boldsymbol{L}}$ is shown in Table 4. Only those parameters that are free to be estimated have t-statistics reported against them. All other parameters are fixed during estimation. It can be seen that there are significant error correlations across different nominal variables and the continuous variable even after including right hand side endogenous variables in the equations that comprise the joint model system. In other words, even after accounting for observed relationships among the tour attributes considered in this paper, there are correlated unobserved factors affecting these attributes leading to the estimation of significant error correlations. The interpretation of the error correlations is that unobserved attributes that affect one dimension are correlated with unobserved attributes that affect another dimension. In this particular study, it is found that all significant error correlations are positive. For example, unobserved factors that contribute to partly joint or joint tours are positively correlated with unobserved factors that contribute to complex tour formation. Suppose a person is a fun-seeking individual who likes to socialize and visit friends. Then, this unobserved attribute of the individual is likely to positively influence both joint tour formation and complex tour formation. Such individuals are likely to enjoy traveling with others (friends) leading to the formation of

[3] The t-statistics reported in the table are with respect to the corresponding values in an independent model where we have 1^s along the diagonal and 0.5^s for all off-diagonal elements in each of the block diagonal matrices corresponding to each nominal variable and 0^s for rest of the elements. It can also be seen that parameters which are fixed during the estimation process do not have t-statistics reported along with them.

joint tours. Such individuals are also likely to visit multiple places and undertake complex tours as they seek to engage in fun activities with friends. They may also have to go to multiple locations to pick up and drop off friends.

Similar interpretations may be applied to other significant error correlations. For example, an adventurous individual may be inclined to undertake complex tours and longer tours in search of destinations that meet the individual's activity preferences. The bottom line is that there are significant error correlations, possibly stemming from attitudes and preferences that make individuals likely to bundle certain choice options together, or built environment and accessibility measures that were not included in the model specifications of this paper. The inclusion of such attributes in the model specifications remains a future research exercise. Unobserved attributes that contribute to an individual choosing the car also positively contribute to the choice of the sports utility vehicle as evidenced by the positive error correlation between auto and SUV vehicle type choices. Unobserved attributes that contribute to joint or complex tour formation are positively correlated with unobserved attributes that contribute to longer tours.

It is interesting to note that there are some key differences in model results between the multi-dimensional choice model system presented in this paper and the bivariate model system estimated on the same data set presented in earlier research (Konduri et al, 2011). In the bivariate model system where accompaniment and complexity were treated as exogenous variables without adequate accounting for endogeneity and correlated unobserved attributes simultaneously impacting these additional dimensions, the tour complexity was found to positively impact choice of van in the vehicle type choice model. However, in the model estimated for this paper, tour complexity was not statistically significant at all in any of the vehicle type choice utility equations. Also, in the previous research effort, the influence of accompaniment on tour complexity was never captured because these two variables were treated as independent variables. In the earlier bivariate model, the number of error correlations that could be estimated was considerably smaller because only two choice dimensions were considered as endogenous. Among the error correlations, only the one between van type choice and tour length was found to be statistically significant. Other relevant error correlations that were found to be significant in this work were not found statistically significant in that simpler bivariate model system. Moreover, the error correlation between van and tour length was found to be negative in that earlier bivariate model. In the multi-dimensional model of this paper, this error correlation is found to be positive, suggesting that model parameter estimates and inferences are significantly impacted by the lack of proper accounting for endogeneity in multivariate modeling contexts. The finding in this paper suggests that unobserved attributes contributing to longer tours (such as living in suburban locations with lower accessibility to destinations) also contribute to the choice of van as a vehicle type (as households in these locations tend to have larger household sizes with children and may desire to use the van to accommodate multiple individuals more comfortably).

Finally, if one were to compare the model estimation results against the original hypothesized structure of the nature of the relationships among these endogenous variables as depicted in Figure 1, it is seen that the relationships postulated in that figure are all significant except the one where tour complexity affects vehicle type choice. It appears that tour complexity itself does not directly affect vehicle type choice. Rather there are common unobserved attributes that simultaneously impact tour complexity and vehicle type choice as evidenced by the positive significant error correlation between tour complexity and car vehicle type choice. However, this covariance is rather weakly significant (with a t-statistic of 1.69) and no other tour complexity – vehicle type choice error covariance is significant. In other words, according to the model estimated in this paper, the relationship between tour complexity and vehicle type choice is quite tenuous, a finding substantially different from the earlier paper (Konduri et al, 2011) where tour complexity was found to significantly directly impact (positively) the choice of van. However, one of the key similarities between the findings of the two studies is that, in both cases, model specifications where vehicle type choice significantly affected tour length were statistically superior to model specifications where tour length affected vehicle type choice. Thus, the notion that vehicle type choice is a longer term decision, where vehicles are broadly allocated to adults or drivers in a household as a higher level household decision, appears to hold true regardless of whether one considers accompaniment and tour complexity as exogenous to the system or endogenous to the system. However, the multidimensional choice model estimation results in this paper point to the influence that accompaniment and tour complexity have on vehicle type choice in the context of a tour. In other words, although vehicle allocation to adults may be occurring as a longer-term higher-level decision process, conscious decisions regarding vehicle type choice and tour length are being made at the tour level depending on the nature of the tour (in terms of accompaniment and complexity).

CONCLUSIONS

This paper presents a multi-dimensional choice model system of tour attributes with a view to better understand the complex inter-relationships that exist among various choice dimensions of interest in the context of tour- and activity-based travel model specification. The four dimensions considered in this paper are passenger accompaniment, tour complexity (measured by number of stops undertaken), vehicle type chosen, and tour length (distance). Modeling these choice dimensions independently of one another, without recognizing the potential presence of correlated unobserved attributes that simultaneously impact multiple dimensions, leads to a number of limitations that may result in erroneous behavioral inferences and travel forecasts. First, when endogeneity exists among multiple choice dimensions that are modeled independently of one another in a series of sequential models loosely strung together, resulting parameter estimates are biased and inconsistent. This can lead to erroneous impact assessments and scenario forecasts. Second, it is entirely possible that some choice dimensions are made as a package or bundle by individuals. In the context of a tour, it is conceivable that choices regarding passenger accompaniment, stop formation, vehicle type, and locations to be visited constitute a package of choices that are made together in a bundle. When that happens, there are bound to be unobserved attitudinal and lifestyle preference variables that inevitably impact multiple dimensions simultaneously. Thus, a model that ignores the bundling or packaging of choices will inevitably be limited in its representation of behavioral processes at play.

This paper makes two major contributions to the field. First, the paper presents an econometric methodology for estimating multi-dimensional choice model systems that include a variety of dependent variable types and accommodate error covariances across multiple dimensions. The modeling methodology takes advantage of the latest advances in model formulation and estimation, and involves the use of novel estimation techniques that greatly reduce computational burden without compromising the efficiency (precision) of parameter estimates. Second, the paper sheds considerable light on the nature of the empirical relationships among the four dimensions examined in this paper. There is much interest in understanding how multiple tour attributes are related to one another with a view to better inform the structure and specification of tour-based models. In addition, there is very limited evidence on tour-level vehicle type choice processes despite the obvious importance of this choice dimension in the ongoing debate regarding energy sustainability and greenhouse gas emission reduction.

A simultaneous equations model is estimated in this paper on a sample of over 6000 tours drawn from the 2009 National Household Travel Survey of the United States. It is found that vehicle type choice is highly dependent on vehicle availability by type, underscoring the need to consider variable choice set composition explicitly when modeling vehicle type choice. The model estimation results show that the dimensions considered in this paper are all related to one another. Passenger accompaniment affects tour complexity, with tours involving passengers likely to be of less complexity involving just a single stop as opposed to multiple stops. Passenger accompaniment, but not tour complexity, affects vehicle type choice with joint tours most likely to be undertaken by van. Passenger accompaniment, tour complexity, and vehicle type choice are all found to affect tour length. Joint tours tend to be longer in distance, as do complex tours involving multiple stops. Van tours tend to be longest in length, followed by car tours. Tours by SUV and pick-up truck tend to be shorter in length than van and car tours. In other words it appears that tours undertaken by more fuel efficient vehicles are likely to be longer than tours undertaken by SUV and pick-up trucks. The results point to the possible conscious choices and decision on the part of travelers to choose locations and travel distances consistent with the fuel efficiency of the vehicle that they drive.

The model in which vehicle type choice affected tour length was found to offer superior statistical fit than the model in which tour length was allowed to affect vehicle type choice. Moreover, the statistical fit of the simultaneous equations model with error covariances was considerably superior to the fit of the independent equations model with error covariances restricted to zero. This finding suggests that there are correlated unobserved attributes simultaneously impact multiple tour dimensions calling for the increased deployment of models such as that presented in this paper. Further research is needed to fully understand the nature of the unobserved attributes affecting these multiple tour dimensions, but these are likely to be personal attitudes and preferences, built environment attributes, and accessibility measures, besides other unobserved variables (such as time-space constraints, household constraints, personal constraints, and institutional constraints) not available in the data set. The modeling methodology presented in this paper has the potential to offer dramatic breakthroughs in the ability of the profession to better capture and represent simultaneous choice processes at play.

From a policy perspective, the findings of the paper suggest that the complex inter-relationships among tour choice dimensions make the analysis of policy impacts potentially more involved than one might have imagined. The findings suggest that the use of a more fuel efficient vehicle for a tour contributes to the choice of a longer tour length. In other words, although the driving of a fuel efficient vehicle may reduce energy consumption and emissions, the finding that it is driven longer distances suggests that the energy consumption and emissions reductions may not be as much as expected and the increase in vehicle miles of travel may actually contribute to greater levels of congestion on roadways. Policies aimed at encouraging the ownership and use of fuel efficient and clean vehicles may end up not providing the originally intended benefits. Another interesting finding is that the flexibility associated with work start time is contributing to the formation of single stop tours (less complexity) of shorter length. In other words, the loosening of time-space constraints imposed by rigid work schedules makes it possible for people to undertake less efficient activity-travel patterns that are characterized by a higher frequency of single stop tours. While an individual single stop tour is likely to be of shorter length than a complex tour, the fact that there are more of them (assuming no change in activity agenda itself) could result in an increase in overall mileage. Again, from a policy perspective, the potential benefits that would be expected from the implementation of a flexible work hours strategy may not be fully realized.

In summary, the paper points to the need to further develop multi-dimensional choice models capable of reflecting the complex observed and unobserved inter-relationships among several behavioral dimensions of interest. Such models would be able to more accurately capture behavioral processes at play and offer more robust forecasts of possible consequences of policy actions. Although the econometric model system formulated and presented in this paper may appear to be a rather complex statistical exercise, it offers the potential to move the profession a step closer to implementing more simultaneous equations model systems that recognize the package nature of activity-travel choices.

REFERENCES

Bagley, M.N. and P.L. Mokhtarian (2002) The Impact of Residential Neighborhood Type on Travel Behavior: A Structural Equations Modeling Approach. *The Annals of Regional Science*, 36(2), pp. 279-297.

Bhat, C.R. (1998) A Model of Post Home-Arrival Activity Participation Behavior. *Transportation Research Part B*, 32(6), pp. 387-400.

Bhat, C.R. (2011) The Maximum Approximate Composite Marginal Likelihood (MACML) Estimation of the Multinomial Probit-based Unordered Response Choice Models. *Transportation Research Part B* (forthcoming).

Bhat, C.R. and H. Zhao (2002) The Spatial Analysis of Activity Stop Generation. *Transportation Research Part B*, 36(6), pp. 557-575.

Bhat, C.R. and N. Eluru (2010) Recent Advances in Discrete and Discrete-Continuous Modeling Systems. In S. Nocera (ed) *Feasibility Decisions in Transportation Engineering*, McGraw-Hill, pp. 111-114.

Bhat, C.R. and S. Sen (2006) Household Vehicle Type Holdings and Usage: An Application of the Multiple Discrete-Continuous Extreme Value (MDCEV) Model. *Transportation Research Part B*, 40(1), pp 35-53.

Bhat, C.R. and S. Srinivasan (2005) A Multidimensional Mixed Ordered-Response Model for Analyzing Weekend Activity Participation. *Transportation Research Part B*, 39(3), pp. 255-278.

Bhat, C.R., J.Y. Guo, S. Srinivasan, and A. Sivakumar (2004) Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns. *Transportation Research Record: Journal of the Transportation Research Board*, 1894, pp. 57-66.

Bhat, C.R., N. Eluru, and R.B. Copperman (2008) Flexible Model Structures for Discrete Choice Analysis. In D.A. Hensher and K.J. Button (eds) *Handbook of Transport Modelling*, 2nd edition, Elsevier Science, pp. 75-104.

Bowman, J.L. and M.E. Ben-Akiva (2001) Activity-Based Disaggregate Travel Demand Model System with Activity Schedules. *Transportation Research Part A*, 35(1), pp. 1-28.

Chung, J. and V.R. Rao (2003) A General Choice Model for Bundles with Multiple-Category Products: Application to Market Segmentation and Optimal Pricing for Bundles. *Journal of Marketing Research*, 40(2), pp. 115-130.

Godambe, V. (1960) An Optimum Property of Regular Maximum Likelihood Estimation. *Annals of Mathematical Statistics*, 31, pp. 1208-1211.

Golob, T.F. (2000) A Simultaneous Model of Household Activity Participation and Trip Chain Generation. *Transportation Research Part B,* 34(5), pp. 355-376.

Golob, T.F. (2003) Structural Equation Modeling for Travel Behavior Research. *Transportation Research Part B*, 37(1), pp. 1-25.

Hamed, M.M. and F.L. Mannering (1993) Modeling Traveler's Postwork Activity Involvement: Toward a New Methodology. *Transportation Science*, 27(4), pp. 381-394.

Kitamura, R., T. Yamamoto, S. Fujii, and S. Sampath (1996) A Discrete-Continuous Analysis of Time Allocation to Two Types of Discretionary Activities Which Accounts for Unobserved Heterogeneity. In J.B. Lesort (ed) *Transportation and Traffic Theory*, Pergamon Press, Oxford, UK, pp. 431-453.

Konduri, K.C., X. Ye, B. Sana, and R.M. Pendyala (2011) A Joint Tour-Based Model of Vehicle Type Choice and Tour Length. *Transportation Research Record: Journal of the Transportation Research Board* (forthcoming).

Mannering, F.L. and C. Winston (1985) A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization. *The RAND Journal of Economics*, 16(2), pp. 215-236.

Mannering, F.L. and D.A. Hensher (2007) Discrete/Continuous Econometric Models and Their Application to Transport Analysis. *Transport Reviews*, 7(3), pp. 227-244.

Mohammadian, A. and E.J. Miller (2003) Dynamic Modeling of Household Automobile Transactions. *Transportation Research Record: Journal of the Transportation Research Board*, 1831, pp. 98-105.

Mokhtarian, P.L. and X. Cao (2008) Examining the Impacts of Residential Self-Selection on Travel Behavior: A Focus on Methodologies. *Transportation Research Part B*, 42(3), pp. 204-228.

Paleti, R., R.B. Copperman, and C.R. Bhat (2011) An Empirical Analysis of Children's After School Out-of-Home Activity-Location Engagement Patterns and Time Allocation. *Transportation*, 38(2), pp. 273-304.

Pendyala, R.M. and C.R. Bhat (2004) An Exploration of the Relationship Between Timing and Duration of Maintenance Activities. *Transportation*, 31(4), pp. 429-456.

Shiftan, Y. and J. Suhrbier (2002) The Analysis of Travel and Emission Impacts of Travel Demand Management Strategies Using Activity-Based Models. *Transportation*, 29(2), pp. 145-168.

Vovsha, P. and M. Bradley (2006) Advanced Activity Based Models in Context of Planning Decisions. *Transportation Research Record: Journal of the Transportation Research Board*, 1981, pp. 34-41.

Wen, C-H. and F.S. Koppelman (1999) Integrated Model System of Stop Generation and Tour Formation for Analysis of Activity Patterns. *Transportation Research Record: Journal of the Transportation Research Board*, 1676, pp. 136-144.

Yamamoto, T. and R. Kitamura (1999) An Analysis of Time Allocation to In-Home and Out-of-Home Discretionary Activities Across Working Days and Non-working Days. *Transportation*, 26(2), pp. 231-250.

Yamamoto, T., R. Kitamura, and S. Kimura (1999) Competing Risks Duration Model of Household Vehicle Transactions with Indicators of Explanatory Variables. *Transportation Research Record: Journal of the Transportation Research Board*, 1676, pp. 116-123.

Zhang, J., H.J.P. Timmermans, and A. Borgers (2005) A Model of Household Task Allocation and Time Use. *Transportation Research Part B*, 39(1), pp. 81-95.

Zhao, Y. and H. Joe (2005) Composite Likelihood Estimation in Multivariate Data Analysis. *The Canadian Journal of Statistics*, 33(3), pp. 335-356.

Table 1. Descriptive Statistics

Variable	Statistic
Tour-level	Mean
Number of passengers on the tour	1.7
Number of trips on the tour	2.7
Number of stops	1.7
Travel duration of the tour	37.0 minutes
Travel length of the tour	15.7 miles
Vehicle type chosen (all tours)	
% Auto	41.9
% Van	14.1
% SUV	25.4
% Pick-up Truck	18.6
Vehicle type chosen (all tours; accounting for vehicle availability in fleet)	
% Auto	51.1
% Van	50.2
% SUV	50.9
% Pick-up Truck	32.3
Household-level	
Average household size	3.1
% 1 person household	3.1
% 2 person household	44.5
% 3 person household	18.4
% 4+ person household	34.0
Average household vehicle count	2.8
% 2 vehicle household	49.9
% 3 vehicle household	30.7
% 4 or more vehicle household	19.4
Household Income	
% households with income < $40K	22.0
% households with income ≥40K and < 60K	17.3
% households with income ≥60K and < 100K	27.8
% households with income ≥100K	27.0
Average number of adults	2.3
Ratio of household size to vehicle count	1.17
Ratio of number of children to number of drivers	0.33
% households in non-urban area	32.0
% households that own the housing unit	95.2
Person-level	
Average age	52.2
% people ≥15 and < 25 years	9.0
% people ≥25 and < 45 years	22.3
% people ≥45 and < 65 years	41.6
% people ≥65 years	27.1
% males	46.0
% Hispanic respondents	6.4
% part-time employees	15.3
% full-time employees	29.2

Table 2. Tour Characteristics by Vehicle Type Chosen and Vehicle Fleet

Vehicle Body Type Selected for the Tour	Vehicle Fleet by Body Type	Frequency	Tour Distance	Travel Time	Number of passengers	Number of Stops
Average Tour Distance (Not Considering Vehicle Fleet Composition)						
Car		2716	16.0	37.7	1.6	1.6
Van		911	15.2	37.0	2.1	1.8
SUV		1647	15.4	36.0	1.8	1.7
Pickup		1204	15.6	36.5	1.5	1.6
Average Tour Distance (Considering Vehicle Fleet Composition)						
Car	Car, Pickup	1204	17.1	39.3	1.6	1.7
Car	Car, SUV	767	14.3	36.1	1.5	1.6
Car	Car, SUV, Pickup	196	16.6	36.9	1.6	1.6
Car	Car, Van	392	15.2	36.5	1.7	1.6
Car	Car, Van, Pickup	99	15.8	35.6	1.6	1.5
Car	Car, Van, SUV	47	17.2	41.2	1.5	1.6
Car	Car, Van, SUV, Pickup	11	20.3	42.4	1.6	1.6
Van	Car, Van	450	15.0	37.5	2.1	1.8
Van	Car, Van, Pickup	102	17.0	38.9	2.1	1.9
Van	Car, Van, SUV	50	11.1	28.8	1.8	1.6
Van	Car, Van, SUV, Pickup	12	21.7	43.8	2.3	1.8
Van	Van, Pickup	169	14.4	35.9	2.0	1.8
Van	Van, SUV	100	17.2	39.6	2.1	1.7
Van	Van, SUV, Pickup	28	15.4	31.9	1.9	1.3
SUV	Car, SUV	824	14.1	34.3	1.7	1.6
SUV	Car, SUV, Pickup	241	16.0	37.4	1.7	1.7
SUV	Car, Van, SUV	46	15.4	36.4	1.5	1.9
SUV	Car, Van, SUV, Pickup	17	18.3	40.4	2.1	1.9
SUV	SUV, Pickup	412	17.0	37.4	1.9	1.8
SUV	Van, SUV	76	16.4	40.3	1.7	1.7
SUV	Van, SUV, Pickup	31	17.9	38.4	1.9	1.6
Pickup	Car, Pickup	662	15.4	35.8	1.5	1.6
Pickup	Car, SUV, Pickup	137	16.9	36.9	1.4	1.6
Pickup	Car, Van, Pickup	51	14.0	33.5	1.6	1.4
Pickup	Car, Van, SUV, Pickup	10	15.0	33.6	1.4	1.6
Pickup	SUV, Pickup	221	15.6	37.4	1.5	1.6
Pickup	Van, Pickup	111	15.9	37.5	1.4	1.7
Pickup	Van, SUV, Pickup	12	17.4	61.4	1.3	1.3

All figures are averages except for the frequency column. Distance is in miles and travel time is in minutes.

Table 3. Model Estimation Results

Variable Description	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
Accompaniment (Base Alternative: Solo)	Partly Joint		Joint			
Constant	-0.9168	-12.51	-0.6350	-11.71		
Socio-economic Attributes						
Ratio of household size to number of vehicles	0.3876	8.24	0.4578	9.55		
Ratio of number of children and number of drivers	0.4953	9.84	0.1875	3.78		
Race of household respondent is Caucasian	-0.1687	-3.27				
Gender (Male = 1, Female = 0)	-0.3065	-7.62	-0.2683	-7.80		
Age 18 years and younger (Yes=1, No=0)	-0.3278	-3.65	0.1688	2.27		
Part-time employment indicator (Yes = 1, No = 0)	0.1252	2.66				
Income indicator: 0 - $40K (Yes = 1, No = 0)			0.1839	4.89		
HH in non-urban area (Yes=1, No=0)			0.0688	2.08		
Tour Complexity (Base Alternative: Simple)	Complex					
Constant	-0.0860	-1.09				
Tour Attributes						
Accompaniment type: Partly Joint (Yes=1, No=0)	-0.3091	-1.98				
Accompaniment type: Joint (Yes=1, No=0)	-0.3197	-1.75				
Socio-economic Attributes						
Can set or change work start time (Yes=1, No=0)	-0.1354	-2.56				
Full-time employment indicator (Yes = 1, No = 0)	-0.0879	-2.20				
Race of household respondent is Hispanic	-0.1326	-1.93				
Age 18 years and younger (Yes=1, No=0)	-0.1648	-2.06				
Gender (Male = 1, Female = 0)	-0.1647	-4.18				
Vehicle Type (Base Alternative: Pickup Truck)	Auto		Van		SUV	
Constant	1.0414	24.21	0.6839	8.34	1.0239	23.22
Tour Attributes						
Accompaniment type: Partly Joint (Yes=1, No=0)	-0.5405	-3.31	0.4617	4.61		
Accompaniment type: Joint (Yes=1, No=0)			0.5195	6.18		
Socio-economic Attributes						
Gender (Male = 1, Female = 0)	-1.0920	-22.43	-1.3309	-18.12	-1.1220	-20.44
Age 65 years and older (Yes=1, No=0)			0.1374	1.77		
No. of children in household			0.0823	3.06		
HH in non-urban area (Yes=1, No=0)			-0.1168	-1.74	-0.0970	-2.27
Tour Length						
Constant	1.4364	14.53				
Tour Attributes						
Accompaniment type: Joint (Yes=1, No=0)	0.6098	3.17				
Tour complexity: Complex tour (Yes=1, No=0)	1.4469	7.89				
Vehicle type: Auto (Yes=1, No=0)	0.0872	2.51				
Vehicle type: Van (Yes=1, No=0)	0.1545	2.81				
Vehicle type: SUV (Yes=1, No=0)	0.0454	1.12				
Socio-economic Attributes						
Ratio of household size to number of vehicles	-0.0843	-2.10				
No. of children in household	-0.0575	-3.29				
Gender (Male = 1, Female = 0)	0.1173	3.91				
Education level: Atleast some college (Yes=1, No=0)	0.0469	1.78				
Can set or change work start time (Yes=1, No=0)	-0.0680	-1.79				
HH in non-urban area (Yes=1, No=0)	0.4470	16.81				
Income indicator: 0 - $40K (Yes = 1, No = 0)	-0.0843	-2.62				
Age 65 years or over (Yes=1, No=0)	-0.0703	-2.34				

Table 4. Error Covariance Matrix

Dimension	Partly Joint	Joint	Complex	Auto	Van	SUV	Tour Length
Partly Joint	1						
Joint	0.5	1					
Complex	0.3525 (4.50)	0.1914 (1.95)	1				
Auto	0.3787 (4.52)	0.1233 (3.97)	0.0469 (1.69)	1			
Van	0	0	0	0.5	1		
SUV	0.1679 (3.96)	0.2057 (5.81)	0	0.6896 (2.39)	0.5	1	
Tour Length	0	0.2030 (1.72)	0.4549 (3.10)	0	0.1278 (2.97)	0	0.9999 (15.59)

Values in parentheses are t-statistics. If no t-statistic is provided, it means that the covariance was fixed to the shown value.

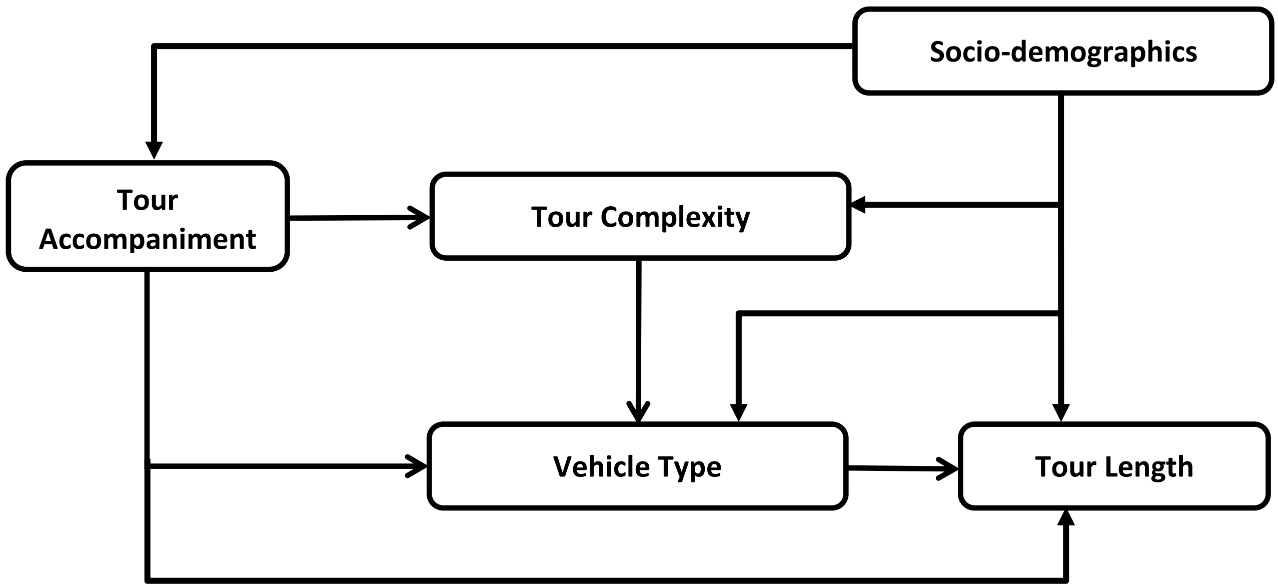

Figure 1. A Framework of Relationships Among Endogenous Variables