Transportation (2009) 36:403–422
DOI 10.1007/s11116-009-9208-x

A copula-based joint multinomial discrete–continuous model of vehicle type choice and miles of travel

Erika Spissu · Abdul Rawoof Pinjari · Ram M. Pendyala · Chandra R. Bhat

Published online: 19 April 2009

Abstract In this paper, a joint model of vehicle type choice and utilization is formulated and estimated on a data set of vehicles drawn from the 2000 San Francisco Bay Area Travel Survey. The joint discrete–continuous model system formulated in this study explicitly accounts for common unobserved factors that may affect the choice and utilization of a certain vehicle type (i.e., self-selection effects). A new copula-based methodology is adopted to facilitate model estimation without imposing restrictive distribution assumptions on the dependency structures between the errors in the discrete and continuous choice components. The copula-based methodology is found to provide statistically superior goodness-of-fit when compared with previous estimation approaches for joint discrete–continuous model systems. The model system, when applied to simulate the impacts of a doubling in fuel price, shows that individuals are more likely to shift vehicle type choices than vehicle usage patterns.

Presented at the 88th Annual Meeting of the Transportation Research Board 11–15 January 2009, Washington, D.C.

E. Spissu
CRiMM – Dipartimento di Ingegneria del Territorio, University of Cagliari, Via San Giorgio 12, 09124 Cagliari, Italy
e-mail: espissu@unica.it

A. R. Pinjari
Department of Civil & Environmental Engineering, University of South Florida, 4202 E. Fowler Avene, ENC 2503, Tampa, FL 33620, USA
e-mail: apinjari@eng.usf.edu

R. M. Pendyala
Department of Civil and Environmental Engineering, Arizona State University, Room ECG252, Tempe, AZ 85287-5306, USA
e-mail: ram.pendyala@asu.edu

C. R. Bhat (✉)
Department of Civil, Architectural & Environmental Engineering, The University of Texas at Austin, 1 University Station, C1761, Austin, TX 78712, USA
e-mail: bhat@mail.utexas.edu

Keywords Vehicle type choice · Vehicle usage · Vehicle miles of travel · Copula-based approach · Discrete–continuous choice modeling · Travel behavior · Greenhouse gas emissions · Transportation energy consumption

Introduction

There is growing consensus in the scientific community that the earth's climate is changing. Global climate change, the broader term used to reflect recent warming trends, has been linked unequivocally to human activity that results in the emission of greenhouse gases. In the United States, energy-related activities account for three-quarters of total human-generated greenhouse gas (GHG) emissions, mostly in the form of carbon dioxide (CO_2) emissions from burning fossil fuels. While about one-half of these emissions come from large stationary sources such as power plants, the transportation sector ranks second and accounts for about one-third of all human generated GHG emissions (EPA 2007). Within the transportation sector, automobiles and light duty trucks (SUVs, pickup trucks, vans and minivans) account for nearly two-thirds of these emissions. Between 1990 and 2003, while emissions from passenger cars increased by just about 2%, GHG emissions from light duty trucks (LDTs) increased by about 50% (EPA 2006). The increase in GHG emissions from automobiles and LDTs reflects the substantial shift in household vehicle fleet composition towards larger, less fuel-efficient vehicles as well as the overall growth in vehicle miles of travel (VMT). The SUV market share, in particular, increased from just about 1% in 1976 to over 25% in 2003, while passenger cars experienced a decrease in share from over 80% to just about 47% during this period (EPA 2006). It is clear that a combination of *vehicle type choice* and *usage (miles traveled)* has contributed to the increase in GHG emissions attributable to the transportation sector.

However, one wonders whether there is a glimmer of hope on the horizon. Over the past 5 years (2003–2008), fuel prices in the United States have increased by as much as 100% (though these fuel prices came down substantially in late 2008). However, it has long been known that travel demand (measured in terms of VMT) is highly inelastic to fuel prices (Hughes et al. 2006; Gicheva et al. 2007). Even with the increase in fuel prices between 2007 and 2008, the decrease in VMT in the United States was only marginal. In fact, the fuel price elasticity of VMT was only of the order of about -0.1. Prior to 2007, VMT continued to rise (albeit at a slower rate) despite increases in fuel prices, suggesting that individuals just absorbed the higher energy costs with virtually no impact on activity-travel demand. The natural next question is: How do household and individual adjust to increases in fuel prices? Recent reports show that households are rapidly moving away from large vehicles in favor of smaller and more fuel-efficient vehicles (Buss 2008). Auto manufacturers are moving forward with the development of alternative fuel vehicles of various kinds. These shifts in consumer demand toward smaller and fuel-efficient vehicles, coupled with new automotive technologies hitting markets around the world, may actually facilitate a continued growth in vehicular travel demand despite the increase in fuel prices.

The above discussion points to the close interplay between vehicle type choice (vehicle fleet composition in households) and usage (vehicle miles of travel) in the transport energy and emission arena. Households adjust to cost structures, socio-economic dynamics, the built environment, and environmental sensitivity by making conscious decisions or choices on the types of vehicles that they will acquire *and* the amount of miles that the vehicles will be driven (Bhat and Sen 2006). In other words, vehicle type choice and usage may be interrelated dimensions of a single choice package rather than two independent choices. These choice dimensions (i.e., type of vehicle and miles of travel) together determine the amount of fuel consumed and the amount of GHG emissions that the household will produce from its travel. It is therefore of interest to model these two choice dimensions *jointly* in an integrated modeling framework.

In this paper, a joint model of household vehicle type choice and usage is formulated and estimated on a data set derived from the 2000 San Francisco Bay Area Travel Survey (BATS). The joint model system recognizes that vehicle type choice and usage are two dimensions of a single choice bundle. That is, the choice of type of vehicle is not an exogenous factor in determining household vehicle miles of travel. On the contrary, vehicle type choice is an endogenous variable in its own right and there may be common unobserved (and, of course, observed) factors that simultaneously influence vehicle type choice and miles of travel. To account for such endogeneity of vehicle type choice, the model takes the form of a joint discrete–continuous structure. The discrete component represents the vehicle type choice dimension and the continuous component represents the miles of travel.

In addition to contributing substantively to the topic of vehicle type choice and usage, the model developed in this paper makes a methodological contribution in the estimation of joint systems with polychotomous (or multinomial) discrete endogenous variables. Most such joint systems have been estimated using either Lee's (1983) full-information maximum likelihood approach or the two-step methods of Hay (1980) and Dubin and McFadden (1984). Lee's approach uses a technique to transform potentially non-normal variables in the discrete and continuous choice equations for each multinomial regime into normal variates, and then adopts a bivariate normal distribution to couple the transformed normal variables. A limitation of Lee's approach is the imposition of a bivariate normal coupling, which allows only linear and symmetric dependencies. The two-step approaches of Hay (1980) and Dubin and McFadden (1984) are based on Heckman's (1974, 1979) method for binary choice situations, and impose a specific form of linearity between the error term in the discrete choice and the continuous outcome (rather than a pre-specified bivariate joint distribution). But these two-step methods do not perform well when there is a high degree of collinearity between the explanatory variables in the choice equation and the continuous outcome equation, as is usually the case in empirical applications, which can lead to unstable and unreliable estimates for the outcome equation (see Leung and Yu 2000, Puhani 2000).

In this paper, we adopt a flexible copula-based approach for estimation of joint discrete–continuous systems with a multinomial discrete choice that generalizes Lee's framework by adopting and testing a whole set of alternative bivariate couplings that can also accommodate non-linear and asymmetric dependencies. Further, the copula approach offers a closed-form expression for evaluating the log-likelihood function in the estimation of model parameters, without requiring any simulation machinery.[1] The Copula approach to discrete–continuous models is based on the concept of a multivariate dependency form (or "copula", which means "link" or "tie" in Latin) for the joint distribution of random variables, in which the dependency is independent of the pre-specified parametric marginal distributions for each random variable (Bhat and Eluru 2009). This concept has been recognized in the statistics field for several decades now, but it is only recently that it has been explicitly recognized and employed in the econometrics field.

[1] In some few cases, simulation-based approaches (such as mixed joint models) that approximate multi-dimensional integrals have also been used to jointly model multinomial discrete choices and continuous outcomes (see, for example, (Pinjari et al. 2007)). However, these approaches involve computationally intensive simulation-based estimation methods.

The remainder of this paper is organized as follows. Following a brief discussion of the literature on modeling vehicle type choice and usage, the paper presents the Copula-based modeling methodology. This is followed by a description of the data and model estimation results. The penultimate section provides results of a policy simulation to demonstrate how the model can be applied to test the impact of changes in fuel prices or any other exogenous factors on household vehicle type choice and usage. The final section offers concluding thoughts and directions for further research.

Modeling vehicle type choice and usage

The analysis and modeling of vehicle type choice and usage has been much of interest to the profession for many years. Several early studies (e.g., Mannering and Winston 1985; Train 1986) examined vehicle type choice in terms of the number of vehicles and vintage. More recent studies, however, have examined vehicle choice in terms of the number of vehicles by type (e.g., Fang 2008; Feng et al. 2005) or vintage and type (Goldberg 1998; Bhat et al. 2009; West 2004). Thus, the focus of research in the vehicle holdings arena has clearly shifted to understanding the type of vehicles possessed by households and this has been largely motivated by energy and environmental concerns, and facilitated by the availability of detailed data about household vehicle holdings. In all these studies, vehicle miles of travel (VMT) serves as the measure of usage.

The studies cited previously employ, for the most part, discrete–continuous model specifications of vehicle ownership (discrete) and utilization (continuous) choices. Typically, the jointness is modeled by capturing the statistical correlation between unobserved variables affecting vehicle type choice and utilization. Many of these studies adopt sequential estimation techniques proposed by Dubin and McFadden (1984) that involve the use of conditional expectation correction terms (West 2004) or instrumental variables (Mannering and Winston 1985; Train 1986; Goldberg 1998). A considerable advance has been made recently in the modeling of vehicle holdings and usage with the development of the multiple discrete–continuous extreme value (MDCEV) model (Bhat and Sen 2006). More recently, Bhat et al. (2009) adopted a joint nested MDCEV-MNL model structure to capture additional dimensions of vehicle holdings.

In contrast to these recent MDCEV-based studies, this paper reverts to the treatment of household vehicle type choice as a simple multinomial choice variable by considering the most recent vehicle purchased by a household. The MDCEV model structure, although very useful to capture the mix of vehicle holdings at any given point in time, fails to capture the dynamics associated with vehicle acquisition. By considering the type of vehicle purchased most recently by a household, one can examine the choice of vehicle type in the context of the other vehicles already owned by the household. Thus, the unit of analysis in this paper is no longer a household as such, but the actual vehicle purchase itself. As in earlier studies, vehicle miles of travel (VMT) is used as the measure of vehicle usage. This leads to the formulation of a more classic joint multinomial logit (MNL)—continuous regression model of vehicle type choice and usage. This formulation constitutes a discrete–continuous model system with the ability to account for endogeneity or self-selection effects (Mannering and Hensher 1987). These effects are captured through error dependencies that account for unobserved factors that affect both vehicle type choice and usage. For example, an individual who “likes to drive” may choose to purchase a certain premium type of car (e.g., high performance car, luxury vehicle) and put many VMT on it. This unobserved personal attribute or preference will then lead to self-selection or error dependency effects. In this way, this paper provides a unique perspective on the dynamics of vehicle purchase decisions as opposed to the MDCEV-based snapshot perspective of household vehicle holdings.[2]

In this paper, we develop a copula-based joint vehicle type choice and usage model to test a host of different dependency surfaces (as opposed to the usual joint normal distribution used de facto in earlier studies) between vehicle type choice and usage equations. The Copula approach to discrete–continuous models is based on the concept of a multivariate dependency form (or "copula") for the joint distribution of random variables, in which the multivariate dependency is independent of the pre-specified parametric marginal distributions for each random variable (Bhat and Eluru 2009). This approach is particularly suited to estimate flexible dependency structures between the discrete vehicle type and continuous usage equations, by allowing one to test several different copulas (see Nelsen 2006) for the joint distribution of the error terms in the two equations (as opposed to the usual joint normal distribution used de facto in earlier studies). Specifically, six different types of copulas (Normal, FGM, Frank, Gumbel, Clayton, and Joe) are tested in the current paper to characterize the dependence structure. In addition, the independent form (with no error correlation) is tested as well. In short, this paper is intended to offer a model capable of determining the extent to which differences in the VMT between different vehicle types are due to "true" effects of vehicle type attributes and policy variables (such as fuel prices), or due to individuals self-selecting to choose vehicle types based on their attitudes, preferences, needs, and desires; and this is done using a novel methodology that obviates the need for adopting less flexible and restrictive model specifications of the past.

Modeling methodology

In this section, we present the structure of the copula-based joint multinomial logit-regression modeling framework to jointly model vehicle type choice and usage. First, the structure of the vehicle type choice model component is discussed, then the vehicle mileage model component is presented, and finally the joint structure between these two model components is described. The procedure used for model estimation is also presented in this section.

The vehicle type choice model component

Let $q(q = 1, 2, \ldots, Q)$ and $i(i = 1, 2, \ldots, I)$ be the indices to represent households and vehicle types, respectively. The vehicle type choice model component takes the familiar discrete choice formulation. Consider the following equation that represents the utility structure of the vehicle type choice model:

$$u_{qi}^{*} = \beta_i' x_{qi} + \varepsilon_{qi} \tag{1}$$

[2] There are two other limitations of the MDCEV approach relative to the more classic discrete-continuous approaches. First, the MDCEV approach ties the discrete and continuous choices in a restrictive framework by having a single stochastic utility function (and therefore, a single error term) that underlies both the discrete and continuous choices. Second, the MDCEV approach needs to have an exogenous total mileage budget of households for implementation.

In the equation above, u^*_{qi} is the latent utility that the qth household derives from acquiring a vehicle of type i, x_{qi} is a column vector of household attributes (including a constant, demographics, and activity-travel environment characteristics) affecting the utility, β_i is the corresponding coefficient (column) vector, and ε_{qi} is the error term capturing the effects of unobserved factors on the utility associated with vehicle type i. With this utility specification, as with any discrete choice model, a household (q) is assumed to choose a vehicle of type i if it is associated with the maximum utility among all I vehicle types; that is, if

$$u^*_{qi} > \max_{j=1,2,\ldots,I,\, j\neq i} u^*_{qj}. \tag{2}$$

Next, following Lee (1983), the polychotomous discrete choice model is recast in the form of a series of binary choice formulations, one for each vehicle type. To do so, let R_{qi} be a dichotomous variable that takes the values 0 and 1, with $R_{qi} = 1$ if the ith alternative is chosen by the qth household and $R_{qi} = 0$ otherwise. Subsequently, substituting $\beta'_i x_{qi} + \varepsilon_{qi}$ for u^*_{qi} [from Eq. 1] in Eq. 2, one can represent the discrete choice model formulation equivalently as:

$$R_{qi} = 1 \quad \text{if} \quad \beta'_i x_{qi} > v_{qi}, \ (i = 1, 2, \ldots, I) \tag{3}$$

$$\text{where} \quad v_{qi} = \left\{ \max_{j=1,2,\ldots,I,\, j\neq i} u^*_{qj} \right\} - \varepsilon_{qi} \tag{4}$$

Equation 3 represents a series of binary choice formulations, which is equivalent to the multinomial discrete choice model of vehicle type. In this equation, the distribution of the v_{qi} term depends on the distributional assumptions of the ε_{qi} terms [see Eq. 4]. The distribution of the v_{qi} terms, in turn, will determine the form of vehicle type choice probability expressions. For example, type-1 extreme value distributed ε_{qi} terms that are independent (across i) and identically distributed imply a logistic distribution for the v_{qi} terms, and, consequently, the vehicle type choice probability expressions resemble the MNL probabilities.

The vehicle mileage model component

The vehicle mileage model component takes the form of the classic log-linear regression, as shown below:

$$m^*_{qi} = \alpha'_i z_{qi} + \eta_{qi}, \quad m_{qi} = 1\left[R_{qi} = 1\right] m^*_{qi} \tag{5}$$

In the equation above, m^*_{qi} is a latent variable representing the logarithm of household (q)'s annual mileage on the vehicle of type i if the household were to choose that type of vehicle in its recent vehicle acquisition. This latent vehicle usage variable is mapped to observed household attributes and the corresponding attribute effects in the form of column vectors z_{qi} and α'_i, respectively, as well as to unobserved factors through a η_{qi} term. On the right hand side of this equation, the notation $1[R_{qi} = 1]$ represents an indicator function taking the value 1 if household q chooses vehicle type i, and 0 otherwise. That is, m^*_{qi} is observed (in the form of m_{qi}) only if household q is observed to hold a vehicle of type i.

The joint model: a copula-based approach

The specifications of the individual model components discussed in the previous two sections may be brought together in the following equation system:

$$R_{qi} = 1 \quad \text{if} \quad \beta_i' x_{qi} > v_{qi}, \quad (i = 1, 2, \ldots, I)$$
$$m_{qi}^* = \alpha_i' z_{qi} + \eta_{qi}, \quad m_{qi} = 1\left[R_{qi} = 1\right] m_{qi}^*. \tag{6}$$

The linkage between the two equations above, for each vehicle type $i(i = 1, 2, \ldots, I)$, depends on the type and the extent of the dependency between the stochastic terms v_{qi} and η_{qi}. As indicated earlier, in this paper, copula-based methods are used to capture and explore these dependencies (or correlations/linkages/couplings). More specifically, copulas are used to describe the joint distribution of the v_{qi} and η_{qi} terms. In this approach, first, the v_{qi} and η_{qi} terms are transformed into uniform distributions using their inverse cumulative distribution functions. Subsequently, copulas are applied to "couple" the uniformly distributed inverse cumulative distributions into multivariate joint distributions. To see this, let the marginal distributions of v_{qi} and η_{qi} be $F_{vi}(.)$ and $F_{\eta i}(.)$, respectively, and let the joint distribution of v_{qi} and η_{qi} be $F_{vi,\eta i}(.,.)$. Subsequently, consider $F_{vi,\eta i}(y_1, y_2)$, which can be expressed as a joint cumulative probability distribution of uniform [0, 1] marginal variables U_1 and U_2 as below:

$$\begin{aligned} F_{vi,\eta i}(y_1, y_2) &= P(v_{qi} < y_1, \eta_{qi} < y_2) \\ &= P\left(F_{vi}^{-1}(U_1) < y_1, F_{\eta i}^{-1}(U_2) < y_2\right) \\ &= P\left(U_1 < F_{vi}(y_1), U_2 < F_{\eta i}(y_2)\right). \end{aligned} \tag{7}$$

Then, by Sklar's (1973) theorem, the above joint distribution (of uniform marginal variables) can be generated by a function $C_\theta(.,.)$ such that:

$$F_{vi,\eta i}(y_1, y_2) = C_\theta(u_1 = F_{vi}(y_1), u_2 = F_{\eta i}(y_2)) \tag{8}$$

where $C_\theta(.,.)$ is a copula function and θ is a dependency parameter (assumed to be scalar), together characterizing the dependency (or correlations/linkages/couplings) between v_{qi} and η_{qi}. The joint distribution formed in the above-discussed manner is used to derive the joint vehicle type choice and vehicle mileage probabilities and log-likelihood expressions.

Model estimation

The joint model has the following log-likelihood expression for a random sample of Q households ($q = 1, 2, \ldots, Q$):

$$L = \prod_{q=1}^{Q} \left[\prod_{i=1}^{I} \left\{ P\left(m_{qi} | \beta_i' x_{qi} > v_{qi}\right) \times P\left(\beta_i' x_{qi} > v_{qi}\right) \right\}^{R_{qi}} \right]. \tag{9}$$

The conditional distributions in the above expression can be expressed as:

$$\begin{aligned} P\left(m_{qi} | \beta_i' x_{qi} > v_{qi}\right) &= \left[P\left(\beta_i' x_{qi} > v_{qi}\right)\right]^{-1} \times \frac{\partial}{\partial m_{qi}} F_{vi,\eta i}\left(\beta_i' x_{qi}, \frac{m_{qi} - \alpha_i' z_{qi}}{\sigma_{\eta i}}\right) \\ &= \left[P\left(\beta_i' x_{qi} > v_{qi}\right)\right]^{-1} \times \frac{1}{\sigma_{\eta i}} \times \frac{\partial}{\partial t} F_{vi,\eta i}\left(\beta_i' x_{qi}, t\right)\Big|_{t = \frac{m_{qi} - \alpha_i' z_{qi}}{\sigma_{\eta i}}} \\ &= \left[P\left(\beta_i' x_{qi} > v_{qi}\right)\right]^{-1} \times \frac{1}{\sigma_{\eta i}} \times \frac{\partial C_{\theta i}\left(u_{q1}^i, u_{q2}^i\right)}{\partial u_{q2}^i} f_{\eta i}\left(\frac{m_{qi} - \alpha_i' z_{qi}}{\sigma_{\eta i}}\right) \end{aligned} \tag{10}$$

where $C_{\theta i}(.,.)$ is the copula corresponding to $F_{vi,\eta i}\left(u_{q1}^{i}, u_{q2}^{i}\right)$ with $u_{q1}^{i} = F_{vi}\left(\beta_i^{'}x_{qi}\right)$ and $u_{q2}^{i} = F_{\eta i}\left(\frac{m_{qi}-\alpha_i^{'}z_{qi}}{\sigma_{\eta i}}\right)$, $\frac{\partial C_{\theta i}\left(u_{q1}^{i}, u_{q2}^{i}\right)}{\partial u_{q2}^{i}}$ is the partial derivative of the copula with respect to u_{q2}^{i} (see Bhat and Eluru 2009), $f_{\eta i}$ is the probability density function of η_{qi}, and $\sigma_{\eta i}$ is the scale parameter of η_{qi}.

Substitution of the above conditional distribution expression back into Eq. 9 provides the following log-likelihood expression for the joint vehicle type choice and usage model:

$$L = \prod_{q=1}^{Q}\left[\prod_{i=1}^{I}\left\{\frac{1}{\sigma_{\eta i}} \times \frac{\partial C_{\theta i}\left(u_{q1}^{i}, u_{q2}^{i}\right)}{\partial u_{q2}^{i}} f_{\eta i}\left(\frac{m_{qi} - \alpha_i^{\prime} z_{qi}}{\sigma_{\eta i}}\right)\right\}^{R_{qi}}\right]. \tag{11}$$

A particular advantage of the copula-based approach is that, in the above log-likelihood expression, a variety of copula [i.e., $C_{\theta i}(.,.)$] functions can be explored to characterize the dependency between vehicle type choice and usage (see Bhat and Eluru 2009 for a review of alternative copula functions available in the literature), and the copulas (hence, the dependency) can be different for different vehicle types. Another appealing feature is that the dependency characterization does not depend upon, and is not limited by, the marginal distributions of v_{qi} and η_{qi}, even if they are differently distributed. However, to complete the model specification, in this paper, we assume that the ε_{qi} terms (for $i = 1, 2, \ldots, I$) associated with the vehicle type choice model component are independent and identically distributed (IID) type-1 extreme value distributed, and that the η_{qi} terms associated with the switching regressions of the logarithm of vehicle mileage follow a normal distribution centered at zero (and, as indicated earlier, with variance $\sigma_{\eta i}^2$). Given these marginal distributions, the log-likelihood expression in Eq. 11 has a closed form expression for most of the copulas available in the literature and hence obviates the need for numerical/simulation-based estimation.

Data description

The primary data set used for this analysis is derived from the 2000 BATS. This survey was designed and administered by MORPACE International, Inc. (2002) for the Bay Area Metropolitan Transportation Commission. The survey collected information on vehicle fleet composition and 2-day activity travel information for over 15,000 households in the San Francisco Bay Area. To each vehicle record from this data, a host of vehicle attributes (such as cost, internal dimensions, performance characteristics, fuel emissions, and fuel type) obtained from the Consumer Guide (2005) and EPA Fuel Economy Guide (EPA 2005) were appended. In addition, residential built environment attributes were constructed and extracted from several secondary sources of data (land use/demographic coverage data, Census 2000 data, and GIS layers of bicycle and transportation network facilities; see Bhat et al. (2009) for a detailed description of data compilation). Finally, based on the 2-day activity-travel information available in the data, each vehicle was assigned to one person (labeled as the primary driver) in the household who drove the maximum number of miles on the vehicle over the 2-day diary period.

In this study, the logarithm of annual vehicle miles traveled (for each vehicle) serves as the continuous dependent variable. Annual vehicle mileage was computed for each vehicle using the odometer readings recorded at the end of the diary period, reported mileage at the time of vehicle possession, the survey year, and the year of possession. The annual vehicle mileage is then:

$$\text{Annual mileage} = \frac{\text{Mileage recorded at end of survey} - \text{Miles on possession}}{\text{Survey year} - \text{Year of possession}}. \tag{12}$$

A log-sum variable was computed from the MNL model results presented in Bhat et al. (2009) for the choice of vehicle make/model for each vehicle type. This log-sum variable contains information on the vehicle attributes, fuel price, and household characteristics (i.e., household size and income) that affected the choice of vehicle make/model within each vehicle type category.

To capture the dynamics of vehicle type choice and usage, this study focuses on "recently acquired vehicles" by the households in the sample. Thus, only those vehicles that were acquired within the preceding 5 year period of the survey were selected for analysis. Vehicles that were purchased prior to the 5 year span were deliberately excluded from the analysis to avoid the data consistency problem; all attribute data is for the year 2000 and hence it was considered prudent to ensure that only those vehicle acquisitions reasonably close to the year 2000 were included in the analysis.

The final sample for analysis includes 3,770 recent vehicle purchase occasions by households. The vehicle purchase at each occasion was classified into one of six vehicle body types, based on the need for an adequate number of chosen instances for each body type: (1) Compact sedans (including subcompact sedans); (2) large sedans (including mid-size sedans and station wagons); (3) coupes; (4) sport utility vehicles (SUV); (5) pickup trucks; and (6) vans (including minivans). Of all these 3,770 recently acquired vehicles, about one-quarter (24.1%) are compact sedans while 30.9% are larger sedans, and 8.2% are coupes. The SUV, pickup truck, and van categories are associated with smaller, but still substantial, percentages (14.7, 11.6, and 10.5%, respectively) in terms of the share of all acquisitions. More importantly, they are associated with higher average vehicle miles of travel, all in excess of 15,500 miles per year. On the other hand, all of the car categories (sedans and coupe) are associated with mileages that are less than 14,500 miles per year. Thus, it appears that larger vehicles are driven more miles, on average, than smaller vehicles—with subsequent implications for energy consumption and emissions.

Model estimation results

This section presents a description of model estimation results for the copula-based joint model of vehicle type choice and vehicle miles of travel. The empirical analysis involved estimating the joint model with all different copula-based dependency structures as well as the independent structure (i.e., independent models). Six different copulas were explored to estimate the jointness between the vehicle choice component and the usage component for each vehicle type. The six types are Gaussian (same as the Lee 1983 specification), FGM, Frank, Gumbel, Clayton, and Joe (a detailed discussion of the nature of each of these copulas is available in Bhat and Eluru (2009); we are unable to provide such a discussion here due to space considerations).

The maximum likelihood estimation of the sample selection model with different copulas leads to a case of non-nested models. Thus, the traditional likelihood ratio test for comparing alternative model specifications is not applicable in this context. An approach to select among the competing copula-based models is the Bayesian Information Criterion (BIC), which collapses to a comparison of the log-likelihood values across different models if all of the competing models have the same exogenous variables and a single copula dependence parameter θ.

It was found that the best model fit was obtained when the Frank copula was used for the continuous regression model associated with all six vehicle types. The log-likelihood value at convergence for the Frank copula-based model is found to be −9,403.47. The likelihood value at convergence for the independent model structure is −9,774.67, clearly rejecting the hypothesis of independence between the vehicle type choice and vehicle usage equations in favor of the model structure that recognizes error correlations. In addition to the final joint model with Frank copulas and the independence model, we estimated a joint model with Gaussian copulas in which all the copulas were specified to be Gaussian (i.e., equivalent to Lee's model). The log-likelihood at convergence for the Gaussian copula-based model was found to be −9,609.96, a significant improvement over the model based on independence, but significantly worse than the Frank copula-based model fit.

The Frank copula-based model estimation results are shown in Table 1. The first numbered-row in the right block of the table shows the copula dependency parameters (and the *t*-statistics in parentheses beneath the parameters) for each vehicle type. As can be observed, all the dependency parameters are significantly different from zero, indicating a significant magnitude of unobserved factors that affect both vehicle type choice and VMT for each type of vehicle. The corresponding Kendall's measures of dependency[3] are: −0.55 (compact sedans), −0.53 (large sedans), −0.56 (Coupe), −0.52 (SUV), −0.58 (Pickup truck) and −0.54 (Vans). To interpret these dependency parameters, note that Eq. 3 can be written as: $R_{qi} = 1 \quad \text{if} \quad \beta_i' x_{qi} - v_{qi} > 0$, and $R_{qi} = 0 \quad \text{if} \quad \beta_i' x_{qi} - v_{qi} < 0$. The error term v_{qi} enters with a negative sign in the equation. Therefore a negative correlation (or dependency) between this error term and the error term η_{qi} in the vehicle usage equation implies that unobserved factors that increase (decrease) the propensity to choose a vehicle of type i also increase (decrease) the usage of that vehicle type. Similarly, a positive correlation between the v_{qi} and the η_{qi} terms implies that unobserved factors that increase (decrease) the propensity to choose a vehicle of type i also decrease (increase) the usage of that vehicle type. Based on intuitive consideration, one can expect the estimated dependency parameters between the v_{qi} and the η_{qi} terms to be negative, so that the dependency between vehicle type choice and usage is positive. As expected, the dependency parameters suggest that unobserved factors that make a household/individual more(less) inclined to acquire a certain vehicle type also make the individual more(less) inclined to put more miles on that vehicle. The magnitudes of the correlation are slightly higher for the coupe and pick-up truck vehicle types, suggesting that there is a higher level of loyalty associated with these vehicle types. These individuals are likely to be those who enjoy driving and enjoy high-performance vehicles; those who are drawn towards these vehicle types are likely to be those who drive and accumulate more miles more than others.

It is interesting to note that the dependency parameters between the v_{qi} and η_{qi}terms obtained using Gaussian copulas (i.e., the Lee (1983) approach) are positive and significant

[3] Kendall's measure of dependency (τ) transforms the dependency parameter (θ) into a number between −1 and 1 (see Bhat and Eluru 2009). For the Frank copula,

$$\tau = 1 - \frac{4}{\theta}\left[1 - \frac{1}{\theta}\int_{t=0}^{\theta} \frac{t}{e^t - 1} dt\right]$$

and $-1 < \tau < 1$. Independence is attained in Frank's copula as $\theta \to 0$.

for all vehicle types with the exception of vans (Gaussian copula estimates are not shown in tables, but are available from the authors). These positive correlations between the error terms are counter-intuitive (see West 2004 for a similar result obtained using the Lee approach). That is, as discussed in the previous paragraph, the implication from the Gaussian copula is that unobserved factors that increase (decrease) the propensity to purchase a certain vehicle type also decrease (increase) the usage of that vehicle type. Further, as indicated earlier, the statistical fit of the joint model using Gaussian copulas is significantly inferior to that using Frank copulas.

The remainder of this discussion (based on Table 1) is intended to provide a description of the impacts of various exogenous variables on the dependent variables of interest in the context of the Frank copula-based model specification that offered the best fit among all the specifications with different copulas.

The first six numbered-columns of Table 1 present the results of the discrete choice component of the model, while the next six numbered-columns present the linear regressions corresponding to usage. The constants (shown in the second row of the table) appear to suggest that in the 5 year period prior to 2000, households tended to acquire SUVs in preference to other vehicle types and had the lowest preference for the acquisition of vans.

The next few rows correspond to individual demographics (age, gender, and race), household socio-demographics (income, presence of children etc.), land use attributes and transportation network measures. Individual demographic effects include the following: (a) The younger age group (16–35 years) tend to acquire compact sedans in comparison to all other vehicle types, while the middle age group (36–55 years) tend to acquire coupes and vans; (b) Males are more likely than females to acquire large sedans, coupes, SUVs, and pick-up trucks, and least likely to acquire vans; and (c) African-Americans are less likely to acquire pick-up trucks and vans, Hispanics are less likely to acquire large sedans and coupes, and Asians are more prone to acquiring sedans and vans.

Among household socio-demographics, households with high income appear to be more likely to acquire large sedans, coupes, SUVs, and vans and less likely to acquire pick-up trucks. The presence of children is generally associated with a propensity to acquire large sedans, SUVs, and vans. The presence of seniors in the household is associated with the purchase of large sedans and vans, but a lower propensity to acquire SUVs. Larger household sizes are associated with the purchase of vans. All of these findings are consistent with expectations and with the large body of literature that speaks to the types of vehicles that households acquire in the context of their socio-demographic characteristics. Finally, among the household variables, it is interesting to note that the variable representing the number of workers was associated with a negative coefficient on four of the six vehicle types. It is likely that these households have already acquired the vehicles that they need and simply did not need to purchase vehicles (other than specialty vehicles such as compact sedan or pick-up truck) in the 5 year period covered by this data set.

Among the land-use attributes, population density did not show a significant impact on vehicle type choice. However, households residing in high employment density areas were found to be less likely to acquire coupes and pick-up trucks. It is likely that pick-up trucks are more suitable to the rugged terrains of suburban/rural areas or the occupational and family needs of households residing in such areas. The land use mix variable provides a rather similar indication. However, it is not immediately clear why the coupe vehicle type also has a negative coefficient associated with its acquisition. The built environment influences may need to be investigated more closely, particularly because the built environment may be endogenous, at least in the long term. As the commercial and industrial

Table 1 Estimation results of the joint model with Frank copulas-parameters (and *t*-statistic)

Variable	MNL						Regression (dependent variable = Log VMT)					
	Compact sedan	Large sedan	Coupe	SUV	Pickup truck	Van	Compact sedan	Large sedan	Coupe	SUV	Pickup truck	Van
Copula dependency parameter (θ)	–	–	–	–	–	–	−6.730 (−12.47)	−6.405 (−12.78)	−6.967 (−8.07)	−6.209 (−9.42)	−7.352 (−10.54)	−6.605 (−8.96)
Constant	–	−0.501 (−2.93)	−0.535 (−3.21)	0.122 (0.74)	−0.433 (−2.30)	−2.396 (−9.91)	8.467 (127.99)	8.610 (183.36)	8.241 (61.47)	8.569 (104.90)	8.437 (76.09)	8.755 (126.34)
Age (age $\geq$ 56 years is base)												
Age between 16 and 35 years	–	−0.362 (−4.52)	−0.362 (−4.52)	−0.362 (−4.52)	−0.362 (−4.52)	−0.362 (−4.52)	0.161 (6.45)	0.161 (6.45)	–	0.161 (6.45)	0.161 (6.45)	0.161 (6.45)
Age between 36 and 55 years	–	–	0.368 (4.23)	–	–	0.368 (4.23)	–	–	–	–	–	–
Male	–	0.142 (1.91)	0.142 (1.91)	0.142 (1.91)	1.564 (11.76)	−0.157 (−1.28)	0.040 (1.92)	0.040 (1.92)	−0.136 (−1.68)	0.040 (1.92)	–	0.040 (1.92)
Ethnicity (Caucasian is base)												
African-American	–	–	–	–	−0.731 (−2.26)	−0.731 (−2.26)	0.261 (3.58)	0.261 (3.58)	0.261 (3.58)	–	–	–
Hispanic		−0.310 (−1.98)	−0.310 (−1.98)	–	–	–	–	–	–	–	–	–
Asian	–	–	−0.346 (−2.54)	−0.346 (−2.54)	−1.508 (−5.02)	–	−0.104 (−2.74)	−0.104 (−2.74)	–	–	–	−0.104 (−2.74)
Other	–	–	–	−0.582 (−1.91)	−0.582 (−1.91)	–	–	–	–	–	–	–
Annual household income												
Annual income (35–90 K)	–	0.350 (2.76)	0.350 (2.76)	0.350 (2.76)	–	0.350 (2.76)	0.084 (3.23)	0.084 (3.23)	–	–	–	–
Annual income (>90 K)	–	0.649 (4.57)	0.649 (4.57)	0.649 (4.57)	−0.205 (−1.68)	0.649 (4.57)	–	–	0.145 (3.41)	0.145 (3.41)	–	–

Table 1 continued

Variable	MNL						Regression (dependent variable = Log VMT)					
	Compact sedan	Large sedan	Coupe	SUV	Pickup truck	Van	Compact sedan	Large sedan	Coupe	SUV	Pickup truck	Van
No. of children in the household												
No. of children ≤ 4 years	–	–	−0.311 (−2.64)	0.246 (3.34)	−0.311 (−2.64)	0.246 (3.34)	−0.053 (−1.75)	–	–	–	0.252 (2.31)	−0.053 (−1.75)
No. of children 5–10 years	–	0.300 (3.87)	–	0.300 (3.87)	0.300 (3.87)	0.300 (3.87)	–	–	0.117 (1.74)	–	0.117 (1.74)	–
No. of children 11–15 years	–	–	−0.591 (−3.11)	–	–	–	–	0.102 (3.51)	–	0.102 (3.51)	–	–
No. of children 16 and 17 years	–	–	–	–	–	–	–	0.081 (1.21)	–	–	–	–
Number of senior adults (>65 years) in the household	–	0.331 (4.88)	–	−0.453 (−3.33)	–	0.331 (4.88)	−0.141 (−6.75)	−0.141 (−6.75)	−0.141 (−6.75)	−0.141 (−6.75)	−0.141 (−6.75)	−0.141 (−6.75)
Household size	–	–	–	–	–	0.495 (11.12)	–	–	−0.069 (−2.21)	–	−0.069 (−2.21)	–
Number of employed individuals in the household	–	−0.163 (−3.67)	−0.163 (−3.67)	−0.163 (−3.67)	–	−0.163 (−3.67)	0.070 (4.45)	0.070 (4.45)	–	–	0.070 (4.45)	–
Land use variables												
Population density	–	–	–	–	–	–	−0.004 (−2.50)	–	–	−0.002 (−1.28)	–	–
Employment density	–	–	−0.002 (−1.99)	–	−0.002 (−1.99)	–	–	–	–	–	–	–
Land use mix (range 0–1)	–	–	−0.544 (−2.70)	–	−0.544 (−2.70)	–	–	–	–	–	–	–

Table 1 continued

Variable	MNL						Regression (dependent variable = Log VMT)					
	Compact sedan	Large sedan	Coupe	SUV	Pickup truck	Van	Compact sedan	Large sedan	Coupe	SUV	Pickup truck	Van
Presence of 4+ physical activity centers	–	–	–	–	–	–	−0.050 (−1.26)	–	−0.050 (−1.26)	−0.050 (−1.26)	–	–
Commercial/ industrial acres within 1 mile radius	–	–	–	−0.001 (−2.31)	–	−0.001 (−1.98)	–	–	–	–	–	–
Local transportation network measures												
Walk access time to in-zone transit stop	–	0.012 (2.04)	–	–	0.012 (2.04)	0.012 (2.04)	–	–	–	–	–	–
No. of zones accessible by bike within 6 miles	–	–	–	–	−0.010 (−4.58)	–	–	–	–	–	−0.006 (−3.88)	–
Presence of old vehicles (same type)	–	0.110 (1.40)	0.110 (1.40)	–	−0.404 (−2.66)	−0.404 (−2.66)	–	–	–	–	–	–
Log-sum parameter	1.000	1.000	1.000	1.000	1.000	1.000	–	–	–	–	–	–
Scale parameter	–	–	–	–	–	–	0.848 (30.99)	0.743 (33.13)	1.047 (19.37)	0.803 (22.99)	1.114 (24.40)	0.765 (22.46)

acreage within a 1 mile radius of the household location increases, the probability of purchasing a SUV or van decreases. This is consistent with the notion that SUVs and vans tend to be vehicles acquired by suburban/rural households that are likely to be farther away from commercial and industrial property.

The transportation network attribute impacts suggest that those who reside in neighborhoods with shorter walk access to transit stops are found to be less likely than those residing in neighborhoods with longer walk access to acquire larger vehicles (large sedans, pick-up trucks, and vans). It is possible that households with short walk access to transit are residing in higher density areas with limited parking space and maneuverability. Hence there is a lower likelihood of acquiring large vehicles. This is further confirmed with the finding that, as the number of zones accessible by bicycle within six miles (or zonal bicycle network connectivity) increases (i.e., as zonal density increases[4]), the probability of purchasing pick-up trucks decreases.

Finally, there is history dependency in vehicle acquisition. If a household already owns a pick-up truck or a van in its fleet, then it is less likely that the household will acquire another one of these vehicle types. On the other hand, if a household already owns a large sedan or a coupe, then the household is more likely to acquire the same vehicle type again. It is conceivable that pick-up trucks and vans are specialty vehicles (large vehicles) and most households do not need more than one of these types of vehicles. Therefore, if one of these vehicle types already exists in the fleet, then the household is unlikely to acquire another one of these. On the other hand, sedans and coupes constitute general purpose automobiles and households may have multiple vehicles of these types for various members of the household.

The log-sum parameter was not found to be statistically different from one, and so is set to one, indicating independence among the utilities of make/model alternatives within each vehicle body type category in vehicle make/model decisions. The corresponding log-sum variable captures the utility derived from the different make/model combinations within each vehicle type.

The second set of six columns includes the linear regressions for the vehicle usage variable. There is one equation for each vehicle type. It is found that young individuals are more likely to drive more than other age groups. Males drive more miles on most vehicle types, except for coupes and pick-up trucks. These findings are rather surprising as one would expect males to put more miles on coupes and pick-up trucks. However, the fact that males are more likely to purchase one of these vehicle types does not necessarily mean that they are going to put more miles on it. Asians are associated with lower mileage on compact and large sedans and vans. African-Americans put more miles on all of the car types—compact and large sedans, and coupes.

Those in the middle income range put more miles on cars, while those in the higher income group accumulate more miles on coupes and SUVs. Those with young children (less than or equal to 4 years of age) put less miles on compact sedans and vans, presumably because of the constraints associated with traveling with very young children. However, as the number of older children increases, households accumulate more miles across a range of vehicle types (as evidenced by the positive coefficients associated with variables representing number of children by age group). Seniors accumulate fewer miles across all vehicle types, larger households put fewer miles on coupes and pick-up trucks,

[4] Zonal bicycle network connectivity represents how small and compact the zones are (i.e., the zonal density).

and households with more workers accumulate more miles on three of the six vehicle types. Virtually all of these findings are consistent with expectations.

Higher population density and the greater presence of physical activity centers in the vicinity of the residential area contribute negatively to the accumulation of miles, particularly for small cars and SUVs. This finding is consistent with the notion that higher densities are associated with lower vehicular miles of travel. Zonal density is also negatively associated with miles accumulated on pick-up trucks.

Finally, the significant scale parameter suggests that there are considerable unobserved factors affecting usage patterns for all vehicle types.

A policy analysis example

This section presents a policy application example using the proposed model system. Specifically, the changes in vehicle type choice and usage are predicted due to an increase in the fuel price from about $2.55 (the fuel price per gallon in the year 2000, converted to current dollars) to $5.00 per gallon; a 96% increase in fuel price. The changes are applied to each vehicle type in the model through the recalculation of the vehicle make/model log-sum variable according to the specification in Bhat et al. (2009). This log-sum variable is used as an explanatory variable in the vehicle type choice model component.

The effect of the fuel price change on aggregate vehicle holdings and usage patterns is measured along two dimensions, i.e., the percent change in acquisition of various vehicle types, and the percent change in the annual vehicle usage (VMT) for each vehicle type. Results of the shifts brought about by the 96% change in fuel costs considered in this study are tabulated in Table 2.[5] The results in Table 2 are presented for three model specifications: (1) The independent model specification; (2) The Gaussian copula-based model specification; and (3) The Frank copula-based model specification. The policy analysis results using the independent model specification suggest a decrease in the market share of SUVs, pickup trucks, and vans, and an increase in the market share of compact and large sedans and coupes (see the first numbered column in the table). Similar results are found using the models with Gaussian copulas and Frank copulas (see the third and fifth numbered columns, respectively). However, notable differences can be found when the vehicle usage changes are compared with the vehicle type market shares across all the three models. *First*, within the results of the independent model (in numbered columns two and three), the percent change in vehicle usage is the same as that for vehicle type choice, reflecting the assumption of independence in this model specification. No jointness is assumed in the model formulation and therefore, the use of each vehicle type simply tracks according to the shift in vehicle type choice. *Second*, the results from the Gaussian copula-based model (i.e., Lee's 1983 model) suggest that the adjustments in vehicle miles of travel will exceed the shifts in vehicle type choice. All of the percent changes in usage are greater than the percent shifts in vehicle type choice (except for vans where it is identical; this is because the corresponding dependency parameter was not statistically different from zero or independence). *Third*, the policy analysis results of the Frank copula-based model suggest the reverse. That is, while there is a shift from larger vehicles to smaller vehicles similar to the indications provided by the Gaussian copula-based model, the magnitude of shift in vehicle usage is smaller than the magnitude of shift in vehicle type choice behavior.

[5] The prediction procedure considers the dependency between the vehicle type and usage equations (see Bhat and Eluru 2009 for details).

Table 2 Impact of increase in fuel price from \$2.55 to \$5.00 per Gallon (96% increase in fuel cost)

	Independent model		Joint model (with Gaussian copulas)		Joint model (with Frank copulas)	
	% Change in holdings of vehicle type	% Change in overall use of vehicle type	% Change in holdings of vehicle type	% Change in overall use of vehicle type	% Change in holdings of vehicle type	% Change in overall use of vehicle type
Compact sedan	1.21	1.21	1.35	1.73	1.25	0.98
Large sedan	0.27	0.27	0.35	0.43	0.28	0.23
Coupe	0.28	0.28	0.37	0.58	0.30	0.26
SUV	−1.56	−1.56	−1.56	−2.14	−1.57	−1.33
Pickup truck	−1.04	−1.04	−1.07	−1.54	−1.04	−0.82
Van	−0.85	−0.85	−0.87	−0.87	−0.88	−0.80
Total	–	−0.11	–	−0.15	–	−0.13

In other words, the Frank copula-based model is suggesting that people will shift vehicle type choices more than they will shift or change vehicle miles of travel (amount of travel undertaken).

The results from the Frank copula-based model are more in agreement with the recent real-world vehicle acquisition and usage trends discussed the “introduction” section that consumers are migrating away from large vehicles to smaller and more fuel-efficient vehicles, but the demand for vehicular travel has remained inelastic to increases in fuel prices. While these simulation results agree with real-world trends, appear intuitive, and corroborate the superiority of the Frank copula-based model over the other models, these results need to be used with caution. This is because the model specification did not directly include the impact of gas prices on vehicle usage equation and neglects the possibility that consumers may be considering trade-offs across different vehicle types not only in the vehicle purchase decisions but also in the vehicle usage decisions.

Conclusions

This paper makes a methodological contribution in the formulation and estimation of discrete–continuous model systems by adopting a copula-based methodology wherein flexible error dependency structures can be accommodated between the discrete and continuous choice equations. To our knowledge, this is the first instance in the econometric literature of the development and application of a copula-based joint model with an endogenous multinomial choice variable rather than a binary choice variable.

The model is applied to jointly estimate and analyze vehicle type choice and usage of recently acquired household vehicles. Model estimation was undertaken on a data set of 3,770 vehicles acquired in a 5 year period just preceding the year 2000 survey of a sample of households in the San Francisco Bay Area. Various copula functions were explored to test the presence of different forms of dependency between vehicle type choice and usage for each vehicle type, and the model with Frank copulas for all vehicle types provided the best statistical fit. The corresponding model estimation results showed the presence of significant unobserved factors contributing to positive dependency between vehicle type choice and usage across all vehicle types. When compared with the results of an independent model (that ignores error correlations) and a Gaussian copula-based model (i.e., the Lee 1983 approach), it was found that the Frank copula-based model offered statistically superior goodness-of-fit. More importantly, when the models were applied in the context of a policy simulation example in which fuel price was increased by 96%, the Frank copula-based model suggested that shifts in vehicle usage are smaller than shifts in vehicle type choice. Given that vehicle miles of travel (VMT) has generally been inelastic to rising fuel prices over the past 5 years, and that vehicle sales figures from automakers show a clear migration of consumers to smaller and more fuel-efficient vehicles, it is likely that the Frank copula-based model offers behaviorally realistic representation of shifts in consumer and travel patterns in response to fuel price hikes.

The model simulation results suggest that habitual behavior or inertial forces play a role in shaping the dynamics of activity-travel patterns of individuals and households (Gärling and Axhausen 2003). While there may be subtle adjustments in activity-travel patterns in response to fuel price shifts (or any other travel demand management strategy), it appears that households may exhibit greater shifts in vehicle type choice with the intent of minimizing the adjustments that need to be made to vehicle miles of travel. The analysis suggests that greater impacts on greenhouse gas emissions and energy consumption may be made by spurring technological innovation, by providing tax incentives for people to shift more quickly to fuel-efficient and low-emission vehicles, and by having automakers (either through voluntary means or through regulatory mechanisms such as raising of corporate average fuel economy or CAFE standards) greatly increase production of smaller fuel-efficient and hybrid-fuel vehicles to meet shifts in consumer demand. Relying on reductions in vehicle miles of travel (VMT) to combat global climate change and dependence on oil may not only prove ineffective, but may also result in degradation of quality of life and slowing of economic activity.

There are at least two important directions for further research. First, future studies would benefit from a better measurement and representation of land-use and transportation network measures in models of vehicle type choice and usage. Second, this study concentrates only on the recently acquired vehicle type and usage. It is possible to get a better picture of the impact of gas prices on vehicular demand for travel when the overall household vehicle usage (across all vehicles rather than just the recently acquired vehicle's usage) is analyzed.

Acknowledgments 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 typesetting and formatting this document.

References

Bhat, C.R., Eluru, N.: A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transp. Res. Part B (2009). doi:10.1016/j.trb.2009.02.001

Bhat, C.R., Sen, S.: Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. Transp. Res. Part B **40**(1), 35–53 (2006)

Bhat, C.R., Sen, S., Eluru, N.: The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. Transp. Res. Part B **43**(1), 1–18 (2009)

Buss, D.: Seismic shift to smaller segments rocks US market, Edmunds' analysis shows. Auto Observer Report. http://www.autoobserver.com/2008/05/seismic-shift-to-smaller-segments-rocks-us-market-edmunds-analysis-shows.html (2008)
Consumer Guide: Consumer guide automotive. http://auto.consumerguide.com/ (2005)
Dubin, J., McFadden, D.: An econometric analysis of residential electric appliance holdings and consumption. Econometrica **52**, 345–362 (1984)
EPA: Fuel economy guide, environmental protection agency and US Department of Energy. Website: http://www.fueleconomy.gov/feg/findacar.htm (2005)
EPA: Greenhouse gas emissions from the U.S. Transportation Sector 1990–2003, EEEOffice of Transportation and Air Quality (6401A) EPA 420 R 06 003 March 2006. www.epa.gov/otaq/climate.htm (2006)
EPA: Climate change: human-related sources and sinks of carbon dioxide. Environmental protection agency and US Department of Energy, 2007. http://www.epa.gov/climatechange/emissions/co2_human.html (2007)
Fang, H.A.: A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. Transp. Res. Part B **42**(9), 736–758 (2008)
Feng, Y., Fullerton, D., Gan, L.: Vehicle choices, miles driven and pollution policies. NBER working paper series, Working Paper 11553, Washington, DC (2005)
Gärling, T., Axhausen, K.W.: Introduction: habitual travel choice. Transportation **30**(1), 1–11 (2003)
Gicheva, D., Hastings, J., Villas-Boas, S.: Revisiting the income effect: gasoline prices and grocery purchases. NBER working paper series, No. 13614, Washington, DC (2007)
Goldberg, P.: The effects of the corporate average fuel efficiency standards in the US. Journal of Industrial Economics **46**(1), 1–33 (1998)
Hay, J.: Selectivity bias in a simultaneous Logit-OLS Model: occupational choice and occupational earnings, O ASH 03150. National Technical Information Service, Rockville (1980)
Heckman, J.: Shadow prices, market wages and labor supply. Econometrica **42**(4), 679–694 (1974)
Heckman, J.: Sample selection bias as a specification error. Econometrica **47**(1), 153–161 (1979)
Hughes, J.E., Knittel, C.R., Sperling, D.: Evidence of a shift in the short-run price elasticity of gasoline demand. The Energy Journal **29**(1), 93–114 (2006)
Lee, L.F.: Generalized econometric models with selectivity. Econometrica **51**, 507–512 (1983)
Leung, S.F., Yu, S.: Collinearity and two-step estimation of sample selection models: problems, origins, and remedies. Comput. Econ. **15**, 173–199 (2000)
Mannering, F.L., Hensher, D.A.: Discrete-continuous econometric models and their application to transport analysis. Transport Reviews **7**(3), 227–244 (1987)
Mannering, F., Winston, C.: A dynamic empirical analysis of household vehicle ownership and utilization. Rand Journal of Economics **16**(2), 215–236 (1985)
MORPACE International, Inc.: Bay Area Travel Survey Final Report. Metropolitan Transportation Commission, CA (2002)
Nelsen, R.B.: An introduction to copulas, 2nd edn. Springer, New York (2006)
Pinjari, A.R., Pendyala, R.M., Bhat, C.R., Waddell, P.A.: Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions, technical paper, Department of Civil, Architectural & Environmental Engineering, The University of Texas at Austin, (2007)
Puhani, P.A.: The Heckman correction for sample selection and its critique. Journal of Economic Surveys **14**, 53–67 (2000)
Sklar, A.: Random variables, joint distribution functions, and copulas. Kybernetika **9**, 449–460 (1973)
Train, K.: Qualitative choice analysis: theory, econometrics, and an application to automobile demand. The MIT Press, Cambridge (1986)
West, S.: Distributional effects of alternative vehicle pollution control policies. Journal of Public Economics **88**, 735–757 (2004)

Author Biographies

Erika Spissu is currently a Research Fellow at the University of Cagliari (Italy). She received her Ph.D. from the University of Palermo and University of Cagliari (Italy) in Transport techniques and economics. She spent the past 2 years at The University of Texas at Austin as a Research Scholar focusing primarily in activity-based travel behavior modeling, time use analysis, and travel demand forecasting.

Abdul Pinjari is an Assistant Professor in the Department of Civil and Environmental Engineering at the University of South Florida, Tampa. His research interests include time-use and travel-behavior analysis,

and activity-based approaches to travel-demand forecasting. He has his Ph.D. from The University of Texas at Austin.

Ram M. Pendyala is a Professor of Transportation Systems in the Department of Civil, Environmental, and Sustainable Engineering at Arizona State University. He teaches and conducts research in travel behavior analysis, travel demand modeling and forecasting, activity-based microsimulation approaches, and time use. He specializes in integrated land use-transport models, transport policy formulation, and public transit planning and design. He is currently the Vice-Chair of the International Association for Travel Behavior Research and is the immediate past chair of the Transportation Research Board Committee on Traveler Behavior and Values. He has his PhD from the University of California at Davis.

Chandra R. Bhat is a Professor in Transportation at The University of Texas at Austin. He has contributed toward the development of advanced econometric techniques for travel behavior analysis, in recognition of which he received the 2004 Walter L. Huber Award and the 2005 James Laurie Prize from the American Society of Civil Engineers (ASCE), and the 2008 Wilbur S. Smith Distinguished Transportation Educator Award from the Institute of Transportation Engineers (ITE). He is the immediate past chair of the Transportation Research Board Committee on Transportation Demand Forecasting and the International Association for Travel Behaviour Research.