

**Household Vehicle Type Holdings and Usage: An Application of the Multiple Discrete-  
Continuous Extreme Value (MDCEV) Model**

**Chandra R. Bhat and Sudeshna Sen**

The University of Texas at Austin, Department of Civil Engineering

1 University Station C1761, Austin, Texas 78712-0278

Tel: 512-471-4535, Fax: 512-475-8744

E-mail: [bhat@mail.utexas.edu](mailto:bhat@mail.utexas.edu), [ssen@mail.utexas.edu](mailto:ssen@mail.utexas.edu)

## **ABSTRACT**

The increasing diversity of vehicle type holdings and the growing usage of vehicles by households have serious policy implications for traffic congestion and air pollution. Consequently, it is important to accurately predict the vehicle holdings of households as well as the vehicle miles of travel by vehicle type to project future traffic congestion and mobile source emission levels. In this paper, we apply a multiple discrete-continuous extreme value model to analyze the holdings and use of multiple vehicle types by households. Data for the analysis is drawn from a 2000 San Francisco Bay Area survey. The model results indicate the important effects of household demographics, residence location variables and vehicle attributes on vehicle type holdings and use. The model developed in the paper can be applied to predict the impact of demographic, land use, and operating cost changes on vehicle type holdings and usage. Such predictions are important at a time when the household demographic characteristics are changing rapidly in the United States. The predictions can also inform the design of proactive land-use, economic, and transportation policies to influence household vehicle holdings and usage in a way that reduces traffic congestion and air quality problems.

## 1. INTRODUCTION

The subject of household vehicle type holdings and use has been the focus of extensive research in the fields of economics, marketing and transportation. There are at least two reasons for this. First, vehicle type holdings and use play a significant role in determining consumer demand for different types of vehicles. Thus, from the perspective of car manufacturers, the preferences for different vehicle types in the overall population, and in demographic subgroups of the population, provide information to design future vehicles, to set production levels of different currently existing vehicle types, and to market vehicles by adopting appropriate positioning and targeting strategies. Second, vehicle holdings and use have an important influence on almost all aspects of the activity and travel behavior of individuals and households. For instance, the 2001 National Household Transportation Survey (NHTS) data shows that 87% of the daily trips in the United States are made by personal-use motorized vehicles, of which almost half are contributed by single-occupant vehicles (see Pucher and Renne, 2003). The increasing usage of motorized personal vehicles, combined with significantly low vehicle occupancy rates, has serious policy implications for traffic congestion and air pollution.

The research in the current paper is motivated from a (public) transportation policy perspective, rather than a car manufacturer perspective, though the results from the research should also serve the purpose of car manufacturers. From a transportation perspective, in addition to the increasing usage of motorized personal vehicles, recent studies suggest an increasing diversity of motorized vehicle type holdings by households. The 2001 NHTS data shows that only about 57% of the personal-use vehicles are cars or station wagons, while 21% are vans or Sports Utility Vehicles (SUV) and 19% are pickup trucks. The increased holdings of vans, SUVs, and pickup trucks, in turn, has led to a surge in the vehicle miles traveled using

these vehicles. This shift from small passenger car vehicle miles of travel to large non-passenger car vehicle miles of travel has implications for roadway capacity, since larger vehicles take up more room on roadways than smaller vehicles. The resulting reduced capacity exacerbates the problem of traffic congestion caused by increasing motorized personal vehicle use. Further, Environmental Protection Agency (EPA) statistics show that an average van, SUV, or pickup truck produces twice the amount of pollutants emitted by an average passenger car. The net result from a traffic management and air quality standpoint is higher traffic congestion levels and more mobile source emissions from the tailpipe of vehicles.

Clearly, it is important to accurately predict the vehicle holdings of households as well as the vehicle miles of travel by vehicle type to support critical transportation infrastructure planning and project mobile source emission levels. The household vehicle-holdings mix and vehicle miles of travel varies depending upon the demographic characteristics of the household, vehicle attributes, fuel costs, travel costs, and the physical environment characteristics (land-use and urban form attributes) of the residential neighborhood. Thus, the substantial changes in the demographic characteristics of households and individuals projected in the next decade and beyond can have a significant impact on household fleet holdings and usage. Similarly, the direct and demographic interaction effects of vehicle attributes, fuel costs, travel costs, and neighborhood characteristics are also likely to impact household fleet holdings and usage. A clear estimate of such impacts will not only help accurate predictions, but can also inform the design of proactive land-use, economic, and transportation policies to influence household vehicle holdings and usage in a way that reduces traffic congestion and air quality problems.

Several earlier studies have examined household vehicle holdings, either in the form of the number and type of vehicles owned, the most recent vehicle purchased, or the type of vehicle

driven most often.<sup>1</sup> The previous studies on household vehicle holdings include the choice of the most recent vehicle purchased or the choice of a new vehicle planned to be purchased (Lave and Train, 1979; Kitamura *et al.*, 2000; Brownstone *et al.*, 2000; Page *et al.*, 2000; Birkeland and Jorgensen, 2001), the make/model/vintage composition of the household vehicle holdings (Manski and Sherman, 1980; Mannering and Winston, 1985), the vehicle which is driven most (Choo and Mokhtarian, 2004), joint choice of vehicle make/model/vintage and vehicle ownership level (Berkovec, 1985, Hensher *et al.*, 1992), joint choice of vehicle make/model/vintage and vehicle acquisition type (Mannering *et al.*, 2002) and joint choice of vehicle type and vehicle age (Berkovec and Rust, 1985; Mohammadian *et al.*, 2003). Choo and Mokhtarian (2004) have provided an excellent review of studies focusing on vehicle type holdings, including details of the dependent variable characterizing vehicle types, the significant explanatory variables used in the analysis, the type of modeling structure applied, and information regarding the data source. The reader will note that some of the studies reviewed in Choo and Mokhtarian (2004) examine aspects of vehicle holding jointly with vehicle usage levels.

The earlier studies discussed above have provided important insights into the factors affecting vehicle type choice and use. All of these studies, to our knowledge, use standard choice models (multinomial logit or nested logit) for the vehicle type dimension and a continuous linear regression model for the vehicle use dimension (if this second dimension is included in the analysis).<sup>2</sup> These earlier studies are able to use standard choice models (where one and only one

---

<sup>1</sup>A number of earlier studies have also focused on vehicle ownership levels and use. These studies are not of immediate interest here since the focus of the current paper is on vehicle type holdings and use. For a comprehensive review of studies on vehicle ownership/use (with no vehicle type holdings analysis), the reader is referred to De Jong *et al.*, 2004.

<sup>2</sup> Some studies, such as Train (1986), use the logarithm of vehicle use as the dependent variable. However, the basic functional relationship between the dependent and independent variables takes a linear regression form.

alternative out of several is selected) because of the way they have framed the dependent variable. Specifically, several studies have examined the vehicle type of the most recent vehicle purchased, or the most driven vehicle, or considered only single-vehicle households. These studies, while useful in limited ways, do not capture the portfolio of vehicle types that a single household may hold at any time (for example, a sedan as well as a minivan). Some other studies have considered multiple vehicle type holdings of a household by treating multiple vehicle choices as if they represented a string of independent (or sequential) single vehicle choice occasions, or by enumerating all the possible combinations of vehicle types as alternatives. The problems associated with these approaches are three-fold. First, these approaches do not recognize that there is intrinsic multiple discreteness in the mix of vehicle types held by households. That is, these studies do not consider that households own a mix of vehicle types to satisfy different functional or variety-seeking needs (such as being able to travel on weekend getaways as a family or to transport goods). Thus, there is diminishing marginal returns (*i.e.*, satiation) in using a single vehicle type, which is the fundamental driving force for households holding multiple vehicle types. Standard discrete choice models are not equipped to handle such diminishing marginal returns or satiation effects.<sup>3</sup> Second, the approach of enumerating all possible combinations of vehicle types can lead to an explosion in the number of alternatives in the choice set. If there are  $J$  vehicle types, the number of alternatives would be  $2^J - 1$ . As an example, if there are five distinct vehicle types, one would have to define 35 alternatives in the standard discrete choice approach. This has the result of leading to a model with several alternative specific variables. Third, modeling the continuous dimension of vehicle use becomes very cumbersome in the above approaches.

---

<sup>3</sup>Standard discrete choice models assume no satiation effects because they use a linear utility specification. That is, the marginal utility of any vehicle type is independent of vehicle usage. The reader is referred to Bhat, 2005 for further details.

In this paper, we apply a multiple discrete-continuous extreme value (MDCEV) model derived from the primitives of utility theory. This model addresses the issue of households potentially holding a mix of different vehicle types, jointly with modeling the annual miles of use of each vehicle type. The MDCEV model was developed recently by Bhat (2005) and is ideally suited for vehicle type and use modeling because it is based on the concept that households hold multiple vehicle types due to diminishing marginal returns from the usage of each vehicle type. From a practical standpoint, the MDCEV model represents a parsimonious model structure. In the current application, we extend the MDCEV model to accommodate unobserved heteroscedasticity and error correlation across the vehicle type utility functions by using a mixing structure, resulting in the mixed MDCEV (or MMDCEV) model.<sup>4</sup>

The rest of this paper is structured as follows. The next section discusses the model structure of the MDCEV and MMDCEV models. Section 3 identifies the data sources, describes the preparation of the data for model estimation, and presents relevant sample characteristics. Section 4 discusses the variables considered in model estimation and the empirical results. Section 5 presents an application of the model. The final section summarizes the major findings of this study and discusses future extensions.

---

<sup>4</sup> The MMDCEV model used here is a static vehicle type holdings and use model, similar to the studies listed earlier in this section. Such static models predict holdings at any particular period without regard to the vehicle holdings in the earlier period. The application of such static models at different and closely-spaced time points can lead to the unrealistic situation of a household holding very different vehicle portfolios between the two time points. However, such static models may be reasonable over longer time periods, as indicated by De Jong *et al.* (2004). Another formulation that seeks to more consistently reflect the actual household vehicle holdings decision process is the dynamic vehicle transaction approach, in which households decide on disposing, adding, replacing, or maintaining the status quo of their current vehicles over time as well as the attributes of any new vehicles entering their household vehicle fleet (see, for example, Hocherman *et al.*, 1983; Gilbert, 1992; HCG, 1995; De Jong, 1996; Bunch *et al.*, 1996). The dynamic approach is particularly appealing for short-to-medium term forecasts. However, such transaction models require a “significant ongoing commitment to collecting panel data” (see Bunch, 2000) and can be relatively cumbersome to apply. Besides, the theoretical linkage between usage and vehicle type holdings is rather tenuous in most dynamic models to date.

## 2. METHODOLOGY

### 2.1 The Multiple Discrete-Continuous Extreme Value (MDCEV) Model

Let there be  $K$  different vehicle types that a household can potentially own. Let  $m_j$  be the annual mileage of use for vehicle type  $j$  ( $j = 1, 2, \dots, K$ ). The utility accrued to a household is specified as the sum of the utilities obtained from using each type of vehicle. Specifically, the utility over the  $K$  vehicle types is defined as:

$$U = \sum_{j=1}^K \psi(x_j)(m_j + \gamma_j)^{\alpha_j}, \quad (1)$$

where  $\psi(x_j)$  is the baseline utility for vehicle type  $j$ , and  $\gamma_j$  and  $\alpha_j$  are parameters (note that  $\psi$  is a function of observed characteristics,  $x_j$ , associated with vehicle type  $j$ ).

As discussed by Kim *et al.* (2002), the utility form in Equation (1) belongs to the family of translated utility functions, with  $\gamma_j$  determining the translation and  $\alpha_j$  influencing the rate of diminishing marginal utility from using a particular vehicle type  $j$ . The function in Equation (1) is a valid utility function if  $\psi(x_j) > 0$  and  $0 < \alpha_j \leq 1$  for all  $j$ . Further, the term  $\gamma_j$  determines if corner solutions are allowed (*i.e.*, a household does not own one or more vehicle types) or if only interior solutions are allowed (*i.e.*, a household is constrained by formulation to own all vehicle types). The latter situation is, of course, of little practical value, since very rarely (if at all) will any household own all different types of vehicles.

The utility form of Equation (1) is flexible enough to accommodate both interior and corner solutions (Kim *et al.*, 2002; Bhat, 2005). In addition, the utility form is also able to accommodate a wide variety of situations characterizing vehicle type preferences based on the values of  $\psi(x_j)$  and  $\alpha_j$  ( $j = 1, 2, \dots, J$ ). A high value of  $\psi(x_j)$  for one vehicle type (relative to

all other vehicle types), combined with a value of  $\alpha_j$  close to 1, implies a high baseline preference and a very low rate of satiation for vehicle type  $j$ . This represents the situation when a household primarily uses only one vehicle type for all its travel needs (*i.e.*, a “homogeneity-seeking” household). On the other hand, about equal values of  $\psi(x_j)$  and small values of  $\alpha_j$  across the various vehicle types  $j$  represents the situation where the household uses multiple vehicle types to satisfy its travel needs (*i.e.*, a “variety-seeking” household). More generally, the utility form allows a variety of situations characterizing a household’s underlying behavioral preferences for different vehicle types.

A statistical model can be developed from the utility structure in Equation (1) by adopting a random utility specification. Specifically, a multiplicative random element is introduced to the baseline utility as follows:

$$\psi(x_j, \varepsilon_j) = \psi(x_j) \cdot e^{\varepsilon_j}, \quad (2)$$

where  $\varepsilon_j$  captures idiosyncratic (unobserved) characteristics that impact the baseline utility for vehicle type  $j$ . The exponential form for the introduction of random utility guarantees the positivity of the baseline utility as long as  $\psi(x_j) > 0$ . To ensure this latter condition,  $\psi(x_j)$  is parameterized further as  $\exp(\beta'x_j)$ , which then leads to the following form for the baseline random utility:

$$\psi(x_j, \varepsilon_j) = \exp(\beta'x_j + \varepsilon_j). \quad (3)$$

The  $x_j$  vector in the above equation includes a constant term reflecting the generic preference in the population toward vehicle type  $j$ . The overall random utility function then takes the following form:

$$U = \sum_j [\exp(\beta' x_j + \varepsilon_j)] \cdot (m_j + \gamma_j)^{\alpha_j} \quad (4)$$

The satiation parameter,  $\alpha_j$ , in the above equation needs to be bounded between 0 and 1, as discussed earlier. To enforce this condition,  $\alpha_j$  is parameterized as  $1/[1 + \exp(-\delta_j)]$ . Further, to allow the satiation parameters to vary across households,  $\delta_j$  is specified as  $\delta_j = \theta_j' y_j$ , where  $y_j$  is a vector of household characteristics impacting satiation for the  $j^{\text{th}}$  alternative, and  $\theta_j$  is a corresponding vector of parameters.

In the current implementation of the model, we assume that the total household annual mileage,  $M$ , accrued across all personal motorized vehicles is known *a priori*<sup>5</sup>. From the analyst's perspective, the individual is then maximizing random utility ( $U$ ) in Equation (4) subject to the constraint that  $\sum_{j=1}^K m_j = M$ , where  $M$  is the total household motorized annual mileage. This constraint implies that the optimal annual miles on only  $K-1$  vehicle types need to be determined, since the annual miles of use for any one vehicle type can be automatically determined from the annual miles of other vehicle types. The implication is that one of the  $K$  vehicle types will have to be considered as the base when introducing a constant or household-specific variables in the utility functions of the  $K$  vehicle types.

The analyst can solve for the optimal usage (say  $m_j^*$ ) of different vehicle types by forming the Lagrangian for Equation (4) with respect to the total miles of travel constraint and applying the Kuhn-Tucker conditions. Assuming that the  $\varepsilon_j$  terms are independently and

---

<sup>5</sup> This is only because we do not have adequate information from the survey to construct a mileage value for use of non-motorized modes of travel. If this information were available, we can add another "vehicle type" category corresponding to non-motorized modes. This category can be considered as an "outside good" which is always "consumed", since households will use non-motorized modes for some amount of their travel (if at least for walking to the personal vehicle). In this instance,  $M$  would correspond to the total annual motorized and non-motorized travel mileage, and the annual motorized mileage would be endogenous to the model. The total annual motorized and non-motorized travel mileage would need to be modeled in an earlier step in this case.

identically distributed across alternatives, and are distributed standard Gumbel, the model simplifies to a remarkably elegant and compact closed form MDCEV structure (see Bhat, 2005 for a derivation). The probability that the household owns  $I$  of the  $K$  vehicle types ( $I \geq 1$ ) is:

$$P(m_i^* > 0 \text{ and } m_s^* = 0; i = 1, 2, \dots, I \text{ and } s = I + 1, \dots, K)$$

$$= \left[ \prod_{i=1}^I c_i \right] \left[ \sum_{i=1}^I \frac{1}{c_i} \right] \left[ \frac{\prod_{i=1}^I e^{V_i}}{\left( \sum_{j=1}^K e^{V_j} \right)^I} \right] (I-1)! , \quad (5)$$

where  $c_i = \left( \frac{1 - \alpha_i}{m_i^* + \gamma_i} \right)$  and  $V_j = \beta' x_j + \ln \alpha_j + (\alpha_j - 1) \ln(m_j^* + \gamma_j)$ .

In the case when  $I = 1$  for a particular household (*i.e.*, only one vehicle type is chosen by the household), the model in Equation (5) collapses to the standard MNL structure. Intuitively, there is no continuous component to be estimated if only one vehicle type is chosen, because the vehicle type chosen will be used for all the travel miles  $M$  of the household, and  $M$  is given as input (all households own at least one vehicle type, since  $M > 0$ ). Thus, the continuous component falls out, and the multiple discrete-continuous model collapses to the MNL structure of single discrete choice with a non linear-in-parameters utility form for  $V_j$  for the household. Note, however, that if each and every household in the sample chooses only one vehicle type, there is no continuous component to estimate and so the  $\alpha_j$ 's have to be constrained to 1, which is what is implicitly done in standard discrete choice models. In this instance, the MDCEV model is identical to the linear-in-parameters MNL model.<sup>6</sup>

---

<sup>6</sup> The MDCEV model formulation also represents a multiple discrete-continuous extension of the single discrete-continuous model formulations of Hausman (1980), Dubin and McFadden (1984), Hanemann (1984), Mannering and Winston (1985), Train (1986), Chiang (1991), Chintagunta (1993), and Arora *et al.* (1998). In the single discrete-continuous models, the discrete alternatives are assumed to be perfect substitutes so that, in the context of the current application, only a single vehicle type is chosen. This may be viewed as a special two alternative case

## 2.2 The Mixed MDCEV (or MMDCEV) Model

The previous section assumed that the  $\varepsilon_j$  terms are independently and identically distributed across vehicle types. However, these assumptions are needlessly restrictive (for example, households who have a predisposition toward an SUV may also be predisposed toward pickup trucks and minivans, since these vehicles allow more passengers to be carried and/or provide more luggage room). Incorporating a more general error structure in the MDCEV model is straightforward through the use of a mixing distribution, which leads to the Mixed MDCEV (or MMDCEV) model. The approach we use in the current paper for the mixing is more straightforward and parsimonious than the one proposed in Bhat (2005). Specifically, the error term,  $\varepsilon_j$ , may be partitioned into two components,  $\zeta_j$  and  $\eta_j$ . The first component,  $\zeta_j$ , is assumed to be independently and identically standard Gumbel distributed across alternatives. The second component,  $\eta_j$ , is allowed to be correlated across alternatives and to have a heteroscedastic scale. Let  $\eta = (\eta_1, \eta_2, \dots, \eta_J)'$ , and assume that  $\eta$  is distributed multivariate normal,  $\eta \sim N(0, \Omega)$ . Then, for given values of the vector  $\eta$ , one can follow the discussion of the earlier section and obtain the usual MDCEV probability that the household holds and uses  $I$  of the  $J$  vehicle types ( $I \geq 1$ ):

---

within the multiple discrete-continuous formulation, with one alternative (say the first) always being consumed. This first alternative may be labeled as “non-motorized travel mode”, an “outside good” which is always “consumed”, since households will use non-motorized modes for some amount of their travel (if at least for walking to the personal vehicle). The second alternative would be a composite of all motorized vehicle types, the baseline utility for which corresponds to the maximum utility across all motorized vehicle types  $j$ . This maximum utility is essentially a log-sum parameter from the standard discrete choice model for vehicle type. With this simplification, the MDCEV model is applicable to the single discrete-continuous case of a single vehicle type choice and corresponding usage. Of course, if the problem at hand is truly a single discrete-continuous one, the customized formulations of Hausman, Dubin and McFadden, Hanemann, and others listed above have the advantage of elegance and simplicity originating from the application of Roy’s identity to an indirect utility function that links the discrete and continuous choices.

$$\begin{aligned}
& P(m_i^* > 0 \text{ and } m_s^* = 0; i = 1, 2, \dots, I \text{ and } s = I + 1, \dots, K) | \eta \\
&= \left[ \prod_{i=1}^I c_i \right] \left[ \sum_{i=1}^I \frac{1}{c_i} \right] \left[ \frac{\prod_{i=1}^I e^{V_i + \eta_i}}{\left( \sum_{j=1}^K e^{V_j + \eta_j} \right)^I} \right] (I-1)!
\end{aligned} \tag{6}$$

The unconditional probability can then be computed as:

$$\begin{aligned}
& P(m_i^* > 0 \text{ and } m_s^* = 0; i = 1, 2, \dots, I \text{ and } s = I + 1, \dots, K) \\
&= \int_{\eta} \left[ \prod_{i=1}^I c_i \right] \left[ \sum_{i=1}^I \frac{1}{c_i} \right] \left[ \frac{\prod_{i=1}^I e^{V_i + \eta_i}}{\left( \sum_{j=1}^K e^{V_j + \eta_j} \right)^I} \right] (I-1)! dF(\eta)
\end{aligned} \tag{7}$$

where  $F$  is the multivariate cumulative normal distribution. The dimensionality of the integration in Equation (7), in the general case, is equal to the number of vehicle types  $K$ .

### 2.3 Estimation of the Mixed MDCEV Model

The parameters to be estimated in the MMDCEV model of Equation (7) include the  $\beta$  vector, the  $\theta_j$  vectors, and the  $\gamma_j$  scalars for each alternative  $j$  (these are embedded in the  $V_j$  values), and the  $\Omega$  variance covariance matrix characterizing the multivariate distribution of the  $\eta$  vector. Let  $\theta$  be a column vector that stacks all the  $\theta_j$  vectors vertically, and let  $\gamma$  be another column vector of the  $\gamma_j$  elements stacked vertically.

We use the maximum likelihood inference approach to estimate the parameters of the MMDCEV model. Introducing the index  $q$  for individuals, we can write the likelihood function as:

$$L(\beta, \theta, \gamma, \Omega) = \sum_{q=1}^Q \log \left[ \int_{\eta_q} \left[ \prod_{i=1}^I c_{qi} \right] \left[ \sum_{i=1}^I \frac{1}{c_{qi}} \right] \left[ \frac{\prod_{i=1}^I e^{V_{qi} + \eta_{qi}}}{\left( \sum_{j=1}^K e^{V_{qj} + \eta_{qj}} \right)^I} \right] (I-1)! dF(\eta_q | \Omega) \right] \quad (8)$$

We apply Quasi-Monte Carlo simulation techniques to approximate the integrals in the likelihood function and maximize the logarithm of the resulting simulated likelihood function across all individuals with respect to  $\beta$ ,  $\theta$ ,  $\gamma$  and  $\Omega$ . In particular, we evaluate the integrand in Equation (8) at different realizations of the  $\eta_q$  draws (for each individual  $q$ ) from a multivariate normal distribution, and compute the average over the different values of the integrand.

In the current paper, we use a Quasi-Monte Carlo (QMC) method to draw realizations for  $\eta_q$  from the multivariate normal distribution. Specifically, we use 200 draws of the Halton sequence (details of the Halton sequence are available in Bhat, 2001; 2003).

One additional issue needs discussion at this point. The Halton draws do not reflect the desired correlation matrix  $\Omega$  of the multivariate distribution of  $\eta_q$ . They are rather univariate draws for each dimension. To translate the univariate Halton draws to the multivariate Halton draws, we apply the Cholesky decomposition of the variance-covariance matrix  $\Omega$  to the univariate draws (see Train, 2003, page 211). In addition, to ensure the positive-definitiveness of the correlation matrix  $\Omega$ , we parameterize the likelihood function in terms of the elements of the Cholesky decomposed matrix of  $\Omega$  rather than using the elements of  $\Omega$  directly. After obtaining the convergent parameter values in terms of Cholesky decomposed-matrix of  $\Omega$ , we obtain the equivalent convergent of values of the elements of the matrix  $\Omega$ .

All estimations were undertaken using the GAUSS programming language. The gradients of the likelihood function (with respect to the parameters to be estimated) were analytically coded for use in the maximum simulated likelihood procedure.

### **3. DATA SOURCES AND SAMPLE FORMULATION**

The primary data source used for this analysis is the 2000 San Francisco Bay Area Travel Survey (BATS). The BATS survey was designed and administered by MORPACE International Inc. for the Bay Area Metropolitan Transportation Commission. The survey collected information on the vehicle ownership of 14529 households in the Bay Area, including the number of vehicles owned by the household, their make and model, year of possession and vehicle usage. The dataset also included information on the sociodemographic, employment and residential location characteristics of these households.

The BATS survey, however, does not include information on fuel economy, fuel use, and fuel cost by vehicle make and model, which are important attributes needed to examine the economic and environmental considerations associated with household vehicle fleet holdings and usage. These fuel-related data were obtained from a secondary data source; the Fuel Economy guide; which is jointly published by the U.S. Environmental Protection Agency and the U.S. Department of Energy (EPA and DOE, 2004).

The sample used in the current analysis included households from the BATS survey with non-zero vehicle ownership. After selecting these non-zero vehicle households, the vehicles owned by each household were categorized into one of five vehicle types based on their make, model and year. The five vehicle types are (1) Passenger car, (2) Sports Utility Vehicle (SUV),

(3) Pickup truck, (4) Minivan, and (5) Van.<sup>7</sup> Some households in the BATS survey did not provide information on vehicle make and model for their vehicle fleet, and these households were removed. From the remaining sample of households, we randomly selected 3500 households for estimation. Of these 3500 households, 1797 (51%) households owned a single vehicle, 1305 (37%) owned two vehicles and the remaining 398 (11%) households owned three or more vehicles.<sup>8</sup>

Table 1 provides information on the distribution of vehicle types in one-vehicle households. This table indicates that most of the one-vehicle households own passenger cars, which include coupes, sedans, hatchbacks and stationwagons. The percentage of one-vehicle households owning SUVs and Pickup trucks is about 11% each. The average annual mileage values (see last column of Table 1) indicate that households owning minivans, SUVs and vans use their vehicles more than households owning passenger cars and pickup trucks.

Table 2 presents information on the distribution of vehicle types within the group of two-vehicle households. The table shows that only 43.1% of the two-vehicle households own both vehicles of the same type. Thus, more than half of the two-vehicle households own vehicles of two different types. The most dominant combination of different vehicle types is the passenger car and pickup truck combination (about 20% of the households), the passenger car and SUV combination (about 16% of the households), and the passenger car and minivan combination (about 12% of the households). Also, the average annual mileage statistics (see the last two columns) shows that households have different usages of different vehicle types. For example,

---

<sup>7</sup> The vehicle type classification used here is oriented toward transportation infrastructure planning and emissions modeling, and hence the detailed make/model of vehicles is not considered. Some earlier studies have used a finer definition of vehicle types to include vehicle makes and models (see Bunch, 2000 for a review).

<sup>8</sup> The sample size of 3500 was based on run-time considerations as well as the judgment that 3500 observations were adequate for accurate and reliable model estimation.

households which own a passenger car and an SUV use their SUV more than the passenger car (see the third row of the Table).

Of the 3500 households in the sample, 326 households (9%) own three vehicles. Within the group of these 326 households, only 18% own vehicles of the same type. About 50% of the households own two passenger cars and a third vehicle of a different type, while 27% own one passenger car and two vehicles of another type.

## **4. EMPIRICAL ANALYSIS**

### **4.1 Variable Specification**

Several different types of variables were considered in the vehicle type and usage model. These included household sociodemographics, residential location variables, and vehicle attributes. The household sociodemographic variables considered in the specifications included household income, presence and number of children, number of employed individuals, presence of disabled individuals and presence of senior adults in the household. The residential location variables included population density of the residential area of the household and the residential area classified into one of four categories: (1) central business district (CBD), (2) urban, (3) suburban and (4) rural area. The only vehicle type attribute we were able to include is the vehicle operating cost per mile, which is defined as the price of a gallon of gas divided by the average vehicle miles per gallon.<sup>9</sup>

---

<sup>9</sup> As per the fuel economy statistics, passenger cars are considered the most fuel efficient vehicles, while pickup trucks and vans are considered the least fuel efficient.

## 4.2 Empirical Results

Table 3 presents the final specification of the model. The final specification was obtained by a systematic process of eliminating insignificant variables and combining the effect of variables, when their impacts were not significantly different. The specification process was also guided by previous literature in the field, and parsimony and intuitive considerations.

### 4.2.1 *Effect of Household Sociodemographics*

Among the set of household sociodemographic variables, the effect of annual household income in Table 3 indicates that high income households are unlikely to own and use pickup trucks and vans. Such households have a higher baseline preference for passenger cars, SUVs and minivans (alternate functional forms for capturing the effect of income were also attempted, but the dummy variable form turned out to be the best in terms of data fit; note, however, that the continuous value of household income appears as a normalization variable to represent the effect of operating costs in section 4.2.3)

The presence of children in the household has a substantial effect on vehicle type choice and use. The results show that households with very small children (less than 4 years of age) have a strong baseline preference for SUVs and minivans, presumably because these vehicles are more spacious, safe, and comfortable for travel with small children. A similar result is found for households with children between 5-15 years of age, except that such households prefer the minivan more and SUV less than households with infants. This result is intuitive, since households with older children have greater space needs and may carpool with other households to transport children. The preference for minivans is strongest among households with young adults.

In addition to the effect of children on the preference for minivans, the results also indicate that households with more individuals prefer minivans to other vehicle types. The preference for minivans, and especially vans, is particularly high for households with one or more mobility challenged individuals, possibly because vans provide ample leg room and are easier to get in and out of.

Finally, the effect of the last two variables under household sociodemographics indicate that households with several employed individuals are not inclined to own and use minivans, while households with many males have a stronger baseline preference for pickup trucks.

#### *4.2.2 Effect of Household Location Variables*

Several household location variables were considered in our specifications, but the only variable that was statistically significant was population density. The results indicate a strong disinclination toward pickup trucks and SUVs among households residing in highly dense neighborhoods. This result deserves further exploration in the future to better understand the nature of this effect. However, one plausible explanation for this effect is that pickup trucks and SUVs are “rugged-terrain” vehicles. Thus, households residing in low density rural areas (which are more likely to be associated with rugged terrains than high density urban areas) are more likely to own and use pickup trucks and SUVs.

#### *4.2.3 Effect of Vehicle Operating Cost*

The only vehicle-type attribute in our analysis is the operating cost for each vehicle type. Our specification tests indicated that it is most appropriate to include this variable relative to the income earnings of the household. As expected, Table 3 indicates that, all other things being

equal, households prefer vehicle types that are less expensive to operate. This effect is particularly pronounced for households with low income.

#### 4.2.4 Baseline Preference Constants

The baseline preference constants do not have any substantive interpretations because of the presence of continuous exogenous variables in the specification. However, since almost all exogenous variables are dummy variables, the constants may be loosely viewed as the generic preference for each vehicle type relative to the base category (*i.e.*, passenger cars). The negative signs on all the constants indicate a general baseline preference for passenger cars relative to other vehicle types.

#### 4.2.5 Satiation Parameters

The satiation parameter,  $\alpha_j$ , for each vehicle type  $j$  is parameterized as  $1/[1+\exp(-\delta_j)]$ , where  $\delta_j = \theta'_j y_j$ . This parameterization allows  $\alpha_j$  to vary based on household sociodemographic and location characteristics, and still be bounded between 0 and 1. In our empirical analysis, no statistically significant variation was found in the  $\alpha_j$  parameters based on household sociodemographics and location.

Table 4a provides the estimated values of  $\alpha_j$  and the t-statistics with respect to the null hypothesis of  $\alpha_j = 1$  (note that standard discrete choice models assume  $\alpha_j = 1$ ). Several important observations may be drawn from the table. First, all the satiation parameters are significantly different from 1, thereby rejecting the linear utility structure employed in standard discrete choice models. That is, there are clear satiation effects in vehicle type usage decisions. Second, satiation effects are low for SUVs and minivans. This perhaps reflects the functionally

versatile nature of these two vehicle types, since they provide comfortable transportation as well as adequate room to carry several people and/or cargo. Hence, households prefer to use these vehicles if they are available to the household. Third, the highest satiation occurs for passenger cars. Of course, passenger cars also have the highest baseline preference compared to other vehicle types. The implication is that households are very likely to own passenger cars, but tend to put more miles on non-passenger car vehicles if such vehicles are available to the household.

#### 4.2.6 Variance-Covariance Parameters

The error components,  $\eta_{qj}$ , introduced in the baseline preference function generate heteroscedasticity and covariance in unobserved factors across the preferences of vehicle types, which is captured by the variance-covariance matrix  $\Omega$  of  $\eta_q$  (See Section 2.2). As indicated in Section 2.2, we do not estimate this variance-covariance matrix directly. Instead, we parameterize the likelihood function in terms of the Cholesky decomposition (say  $S$ ) of  $\Omega$ . After obtaining the estimates of  $S$ , the matrix  $\Omega$  needs to be computed as  $\Omega = S'S$ . The relevant standard errors (and t-statistics) of the elements of  $\Omega$  are computed by re-writing the likelihood directly in terms of  $\Omega$  ( $\Omega$ -parameterized likelihood function), computing the estimate of  $\Omega$  from the estimate of  $S$  at convergence of the  $S$ -parameterized likelihood function, and maximizing the  $\Omega$ -parameterized likelihood function. This “optimization” will immediately converge and provide the necessary standard errors for the elements of  $\Omega$ .

The estimated variance-covariance matrix ( $\hat{\Omega}$ ) is shown in Table 4b. For ease of discussion and because of the symmetric nature of the matrix, only the upper triangle is presented. The reader will note that some of the variance and covariance matrix are zero because they did not turn out to be statistically different. The matrix shows that there is least uncertainty

in the valuation of the passenger car vehicle type relative to other vehicle types (the passenger car uncertainty is confined to the gumbel distributed error term  $\zeta_j$  in section 2.2). The most uncertainty is in the valuation of the van (see the diagonal of the matrix). Further, the results indicate statistically significant covariance in the utilities of the SUV and pickup truck vehicle types, and also the SUV and minivan vehicle types, and the minivan and pickup truck vehicle types. That is, unobserved factors that lead to an increased preference for the SUV also lead to an increased preference for the pickup truck and minivan vehicle types. Similarly, unobserved factors increasing the preference for pickup trucks also increase the preference for minivans.

#### *4.2.7 Overall Measures of Fit*

The log-likelihood value at convergence of the final mixed multiple discrete-continuous extreme value (MMDCEV) model is -9425. The corresponding value for the MMDCEV model with only the constant parameters (in the baseline preference), the satiation parameters, and the variance-covariance terms is -9575. The likelihood ratio test value for testing the presence of exogenous variable effects is 300, which is substantially larger than the critical chi-square value with 14 degrees of freedom at any reasonable level of significance. This clearly indicates variations in the baseline preferences for the vehicle types based on household demographics, household location variables, and vehicle operating costs. Further, the log-likelihood value at convergence of the MDCEV model that does not allow unobserved heteroscedasticity and correlation across the baseline preferences of the different vehicle types is -9445. The likelihood ratio test value for comparing the MDCEV model with the MMDCEV model is 40, which is larger than the critical chi-square value with 7 degrees of freedom (corresponding to the seven parameters estimated to characterize the variance-covariance matrix). Thus, there is statistically

significant unobserved variation across individuals in their baseline preferences, and statistically significant correlation in the utilities of the different vehicle types.

In addition to the likelihood-based measures of fit, one can also obtain more intuitive measures of predictive utility by comparing predicted values of vehicle type ownership and use with the actual observed values at the household-level. The predicted values of vehicle type ownership and use can be obtained by solving the following constrained optimization problem (in the expression below, we use the index  $q$  for households):

$$\text{Max } \tilde{U}_q = \int_{\eta_{q1}=-\infty}^{\infty} \int_{\zeta_{q1}=-\infty}^{\infty} \int_{\eta_{q2}=-\infty}^{\infty} \int_{\zeta_{q2}=-\infty}^{\infty} \cdots \int_{\eta_{qK}=-\infty}^{\infty} \int_{\zeta_{qK}=-\infty}^{\infty} \sum_j \left\{ \left[ \exp(\beta'x_{qj} + \zeta_{qj} + \eta_{qj}) \right] \cdot (m_{qj} + \gamma_j)^{\alpha_{qj}} \right\} \\ dG(\zeta_{q1})dG(\zeta_{q2})\dots dG(\zeta_{qK})dF(\eta|\Omega)$$

subject to

$$\sum_j m_{jq} = M_q, m_{jq} \geq 0 \text{ for all } j,$$

where  $G$  is the standard cumulative Gumbel distribution and  $F$  is the multivariate normal distribution function. The constrained optimization problem above can be solved using simulation techniques.

Summary disaggregate non-likelihood measures of fit can be computed in several ways based on a comparison of actual and predicted values. Two measures of fit are presented here to reflect the discrete as well as continuous nature of the predictions from the MMDCEV model. The first measure evaluates the ability of the model to correctly predict holdings of the various vehicle types (this is the discrete component of the model). This measure, which we label as the “hit rate” measure, indicates the percentage of correct predictions regarding vehicle holdings across all households and vehicle types. This “hit rate” measure is computed to be 84%, a rather respectable value. The second measure evaluates the ability of the model to predict the annual

miles of travel conditional on a correct prediction regarding vehicle holdings (this is the continuous component of the model). This measure, computed as the mean absolute percentage error (MAPE) ratio, is 21%. Overall, the vehicle type model estimated here appears to provide reasonable prediction fits.

## **5. MODEL APPLICATION**

The model estimated in the paper can be used to determine the change in vehicle type holdings and usage due to changes in independent variables over time. This is particularly important because of changing demographic, employment-related and operating cost trends. For instance, the structure of the household is changing rapidly with an increase in households with no children (Texas State Data Center, 2000). The number of employed individuals in the household is also on the rise and this trend is likely to continue despite the short-term slump due to the economy (U.S. Census Bureau, 1999). Such sociodemographic and other changes will have an effect on vehicle type choice and usage, and the model in this paper can be used to assess these impacts.

The prediction method to assess the changes in vehicle type ownership and use in response to changes in relevant exogenous variables, is identical to the one described in the earlier section to obtain the intuitive measures of fit. In this paper, we demonstrate the application of the model by studying the effect of an increase in vehicle operating costs due to an increase in gas cost. Specifically, we modify the “operating cost divided by household income” variable to reflect an increase from the \$1.40 per gallon cost used in estimation (this corresponds to the fuel cost in 2000) to \$2.00 per gallon (the cost in the recent past). To examine the impact

of this increase, we compute revised expected aggregate shares and the total miles of usage of each vehicle type, and then obtain a percentage change from the baseline estimates.<sup>10</sup>

Table 5 presents the results, which show a marginal percentage decrease in the holdings of passenger cars, and more significant decreases in the holdings of all the other vehicle types (see column labeled “percentage change in holdings of vehicle type”). It is interesting to note that the ownership of SUVs and minivans drop by the largest percentage. Though the operating costs of pickup trucks and vans are higher than SUVs and minivans, pickup trucks and vans also have a larger error variance (Table 4b). Consequently, the signal (cost increase) to noise (error variance) ratio is lower for pickup trucks and vans, which has the result of attenuating the impact of the signal (see Bhat, 1995). Intuitively, households who own pickup trucks and vans are more “committed” to these vehicle types than are SUV- and minivan-owning households.

The percentage change in overall usage shows a mild positive increase in the passenger car annual miles of travel, and a higher negative decrease in the annual miles of travel of other vehicle types. This effect combines the holding change effect with the usage change effect. Thus, the overall positive percentage increase in passenger car miles of travel is because of the relatively low drop in passenger car holdings combined with usage switching from non-passenger car vehicle types to the passenger car. Additionally, the positive percentage increase in passenger car miles of travel may also be attributed to  $M$  (the total motorized miles of travel across all vehicle types) being held fixed and exogenous in the current empirical analysis.<sup>11</sup> The overall reduction in usage of the non-passenger vehicle types is consistent with the fact that vans and pickup trucks are the most expensive to operate per mile. It appears that the high operating

---

<sup>10</sup> An increase in vehicle operating costs would likely also impact other travel choices, such as travel mode and destination choice. These impacts are not modeled here.

<sup>11</sup> As indicated in the footnote on page 8, the exogeneity of  $M$  is maintained because of data limitations.

cost signal is strong enough to dominate any differential “noise” effects across the vehicle types when both ownership and usage are considered together.

## 6. CONCLUSIONS

The increasing diversity of vehicle type holdings and the growing usage of non-passenger car vehicles have serious policy implications for traffic congestion and air pollution. Consequently, it is important to accurately predict the vehicle holdings of households as well as the vehicle miles of travel by vehicle type to project future traffic congestion and mobile source emissions levels. The current paper presents the application of a utility-based model for multiple discreteness that models the simultaneous holdings of multiple vehicle types (passenger car, SUV, pickup truck, minivan and van), as well as determines the continuous miles of usage of each vehicle type, in a joint modeling system. The specific model used here is the mixed multiple discrete-continuous extreme value (MMDCEV) model, as recently developed by Bhat (2005). Data for the analysis is drawn from the 2000 San Francisco Bay Survey.

The analysis considered several different kinds of variables to explain vehicle type holdings and usage, including household sociodemographics, household residential location variables and vehicle attributes. Important findings from the analysis are as follows:

1. As the number of children in the household increase, there is a higher preference to own and use SUVs and minivans relative to passenger cars, pickup trucks and vans.
2. Households with several individuals have a higher preference for minivans than households with fewer individuals.
3. Households with one or more mobility-challenged household members are more likely to own and use vans and minivans than households with no mobility-challenged members.

4. Households with more number of employed individuals are less likely to prefer minivans than households with fewer employed individuals.
5. Households with more men in the household prefer pickup trucks to other vehicle types.
6. Households located in densely populated neighborhoods have a disinclination for pickup trucks.
7. Vehicle operating cost has a negative influence on vehicle ownership and usage for all vehicle types except passenger cars.
8. Households are very likely to own passenger cars but put more miles on non-passenger car vehicles if such vehicles are available in the household.

The model estimated in this paper can be used to determine the change in vehicle type holding and usage due to changes in independent variables over time. This is particularly important because of changing demographic, employment-related, and operating cost trends. In the current paper, we demonstrate the value of the model by assessing the impact of an increase in vehicle operating costs, and examining the implications for vehicle type ownership and usage.

To summarize, this paper uses a modeling structure (*i.e.*, the MDCEV structure) that is ideally suited for vehicle type and use analysis because the structure is based on the concept that households hold multiple vehicle types due to diminishing marginal returns from usage of each vehicle type (or, equivalently, due to the need to satisfy different functional or variety-seeking desires of the household). The MDCEV model is a simple and parsimonious model structure, which can be extended to accommodate heteroscedasticity and error correlation across the vehicle type utility function by using an appropriate mixing distribution leading to the mixed

MDCEV (MMDCEV) model. In future research, we plan to embed the MMDCEV model of vehicle type within a larger model system of vehicle type, make, model, vintage, age and usage.

### **ACKNOWLEDGEMENTS**

The authors would like to thank Chuck Purvis of the Metropolitan Transportation Commissions (MTC) in Oakland for providing help with data related issues. The authors also appreciate the valuable comments of an anonymous reviewer on an earlier version of the paper.

**REFERENCES**

- Arora, N., G.M. Allenby, and J.L. Ginter (1998). A hierarchical Bayes model of primary and secondary demand. *Marketing Science*, 17, 29-44.
- Berkovec, J. (1985). Forecasting automobile demand using disaggregate choice models. *Transportation Research Part B*, 19(4), 315–329.
- Berkovec, J. and Rust, J. (1985). A nested logit model of automobile holdings for one vehicle households. *Transportation Research Part B*, 19(4), 275–285.
- Bhat, C.R. (1995). A Heteroscedastic Extreme Value Model of Intercity Mode Choice. *Transportation Research Part B*, 29(6), 471-483.
- Bhat, C.R. (2001). Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B*, 35(7), 677-693.
- Bhat, C.R. (2003). Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. *Transportation Research Part B*, 37(9), 837-855.
- Bhat, C.R. (2005). A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions. In press, *Transportation Research Part B*.
- Birkeland, M.E. and Jordal-Jorgenson, J. (2001). Energy efficiency of passenger cars. Paper presented at the *European Transport Conference 2001*, PTRC, Cambridge, UK.
- Brownstone, D., Bunch, D., and Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles, *Transportation Research Part B: Methodological*, 34(5), pp. 315–338.
- Bunch, D.S., Brownstone, D. and Golob, T.F. (1996). A dynamic forecasting system for vehicle markets with clean-fuel vehicles, *World Transport Research*, 1, pp. 189–203.
- Bunch, D.S. (2000). Automobile demand and type choice, in: D. A. Hensher and K. J. Button (Eds) *Handbook of Transport Modelling*, pp. 463–479 (Oxford: Pergamon).
- Chiang, J. (1991). A simultaneous approach to whether to buy, what to buy, and how much to buy. *Marketing Science*, 4, 297-314.
- Chintagunta, P.K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, 12, 194-208.
- Choo, S. and Mokhtarian, P.L. (2004). What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice. *Transportation Research Part A*, 38(3), 201–222.

- De Jong, G.C. (1996). A disaggregate model system of vehicle holding duration, type choice and use, *Transportation Research Part B: Methodological*, 30(4), pp. 263–276.
- De Jong, G.C., Fox, J., Daly, A., Pieters, M., and Smit, R. (2004) A comparison of car ownership models; *Transport Reviews*, Vol. 24, No. 4, 379-408.
- Dubin, J.A. and D.L. McFadden (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 52(2), 345-362.
- Gilbert, C.S. (1992). A duration model of automobile ownership, *Transportation Research Part B: Methodological*, 26(2), pp. 97–114.
- HCG (1995). Further Development of a Dynamic Vehicle Transactions Model. Report 5037–1 (The Hague: Hague Consulting Group).
- Hannemann, M. (1984). The discrete/continuous model of consumer demand. *Econometrica*, 52, 541-561.
- Hausman, J.A. (1980). The effects of wages, taxes and fixed costs on women's labor force participation; *Journal of Public Economics* 15, 161-194.
- Hensher, D.A., Barnard, P.O., Smith, N.C., and Milthorpe, F.W. (1992). Dimensions of automobile demand, A longitudinal study of automobile ownership and use, North-Holland, Amsterdam.
- Hocherman, I., Prashker, J.N. and Ben-Akiva, M. (1983). Estimation and use of dynamic transaction models of automobile ownership, *Transportation Research Record*, 944, pp. 134–141.
- Kim, J., Allenby, G.M., and Rossi, P.E. (2002). Modeling consumer demand for variety. *Marketing Science*, 21, 229-250.
- Kitamura, R., Golob, T.F., Yamamoto, T., Wu, G. (2000). Accessibility and auto use in a motorized metropolis. Paper presented at the 79th Transportation Research Board Annual Meeting, Washington, DC.
- Lave, C.A. and Train, K. (1979). A disaggregate model of auto-type choice. *Transportation Research Part A*, 13(1), 1–9.
- Mannering, F. and Winston, C. (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics*, 16(2), 215–236.
- Mannering, F., Winston, C., Starkey, W. (2002). An exploratory analysis of automobile leasing by US households. *Journal of Urban Economics*, 52(1), 154–176.

- Manski, C.F. and Sherman, L. (1980). An empirical analysis of household choice among motor vehicles. *Transportation Research Part A*, 14(6), 349–366.
- Mohammadian, A. and Miller, E.J. (2003). An empirical investigation of household vehicle type choice decisions. Forthcoming, *Transportation Research*.
- Page, M., Whelan, G. and Daly, A. (2000). Modelling the factors which influence new car purchasing. Paper presented at the *European Transport Conference 2000, PTRC*, Cambridge, UK.
- Pucher, J. and Renne, J.L. (2003). Socioeconomics of urban travel: Evidence from 2001 NHTS. *Transportation Quarterly*, 57(3), 49-77.
- Train, K. (1986). *Qualitative Choice Analysis: Theory, Econometrics and an Application to Automobile Demand* (Cambridge, MA: MIT Press).
- Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, United Kingdom.
- U.S. Environmental Protection Agency (EPA) and U.S. Department of Energy (DOE) (2004). Fuel Economy guide.

**Table 1. Vehicle Type Distribution of One – Vehicle Households**

<b>Vehicle Type</b>	<b>Total number of households</b>	<b>Percentage of households</b>	<b>Mean Annual Mileage (in miles)</b>
Passenger Car	1298	72%	9299
Sports Utility Vehicle (SUV)	204	11%	10052
Pickup Truck	192	11%	9981
Minivan	89	5%	11193
Van	14	1%	10330

**Table 2. Vehicle Type Distribution Among Two – Vehicle Households**

Type of first vehicle	Type of second vehicle	Total number of households	Percentage total number of two-vehicle household	Mean Annual Mileage of vehicle type 1 (in miles)	Mean Annual Mileage of vehicle type 2 (in miles)
Passenger Car	Passenger Car	522	40.0%	19147*	19147*
Passenger Car	Pickup Truck	255	19.5%	10051	9591
Passenger Car	SUV	213	16.3%	9590	10736
Passenger Car	Minivan	151	11.6%	9841	10171
SUV	Pickup Truck	46	3.5%	9251	10502
Pickup Truck	Minivan	32	2.5%	10514	10524
Pickup Truck	Pickup Truck	21	1.6%	21587*	21587*
SUV	Minivan	17	1.3%	10993	11390
Passenger Car	Van	15	1.1%	7597	9549
SUV	SUV	13	1.0%	24481*	24481*
Minivan	Minivan	7	0.5%	25109*	25109*
SUV	Van	6	0.5%	9736	13564
Pickup Truck	Van	6	0.5%	15172	9509
Minivan	Van	1	0.1%	12014	9455

\* These numbers represents the mean total annual miles across both vehicles. Note that the annual mileage is computed for each vehicle type; in case both vehicles are of the same type, the entries correspond to the total miles across both vehicles. The numbers are the same across the “Mean annual mileage of vehicle type 1” and “Mean annual mileage of vehicle type 2” for this reason.

**Table 3. Empirical Results**

<b>Explanatory variables</b>	<b>Parameter</b>	<b>t-statistic</b>
<b>Household sociodemographics</b>		
<i>Income greater than 115K</i>		
Pickup Truck	-0.6135	-4.683
Van	-0.8684	-1.517
<i>Presence of children less than 4 years of age</i>		
SUV and Minivan	0.6010	3.926
<i>Presence of children between 5 and 15 years of age</i>		
SUV	0.4090	3.836
Minivan	0.7099	4.611
<i>Presence of children between 16 and 17 years of age</i>		
Minivan	0.8416	3.355
<i>Household size</i>		
Minivan	0.5341	5.593
<i>Presence of a mobility-challenged individual in the household</i>		
Minivan	0.3912	1.433
Van	2.1069	1.951
<i>No. of employed persons in the household</i>		
Minivan	-0.3686	-3.775
<i>No. of males</i>		
Pickup Truck	0.3257	4.207
<b>Household location variables</b>		
<i>Population Density</i>		
Pickup Truck and SUV	-0.0166	-4.143
Vehicle Operating cost(cents/mile) divided by household income (in 000s)		
	-0.0314	-2.139
<b>Baseline preference constants</b>		
Passenger Car (base)		
SUV	-3.6514	-11.045
Pickup Truck	-3.1273	-9.199
Minivan	-5.5305	-10.592
Van	-12.5287	-4.584

**Table 4a. Satiation Parameters**

Vehicle Type	Parameter	t-statistic <sup>12</sup>
Passenger Car	0.4410	11.53
Sports Utility Vehicle (SUV)	0.9003	4.90
Pickup Truck	0.7293	6.55
Minivan	0.8480	4.04
Van	0.5124	2.34

**Table 4b. Variance-Covariance Matrix**

Vehicle Type	Vehicle Type				
	Passenger Car	SUV	Pickup Truck	Minivan	Van
Passenger Car	0	0	0	0	0
Sports Utility Vehicle (SUV)		2.32 (4.07)	2.24 (4.48)	1.51 (3.18)	0
Pickup Truck			3.35 (3.10)	1.46 (3.74)	0
Minivan				1.95 (1.98)	0
Van					28.94 (2.47)

<sup>12</sup> The t-statistic is computed for the null hypothesis that the satiation parameter is equal to 1. Equivalently, the t-statistic is for the test that there are no satiation effects or that the utility structure is linear.

**Table 5. Impact of an increase in operating (fuel) cost from \$1.40 per gallon to \$2.00 per gallon**

<b>Vehicle Type</b>	<b>Percentage change in holdings of vehicle type</b>	<b>Percentage change in overall use of vehicle type</b>
Passenger Car	- 0.1	+ 0.5
Sports Utility Vehicle (SUV)	- 5.9	- 3.0
Pickup Truck	- 2.1	- 6.2
Minivan	- 4.9	- 2.3
Van	- 3.4	- 6.5