

## **DEVELOPMENT OF A VEHICLE FLEET COMPOSITION MODEL SYSTEM FOR IMPLEMENTATION IN AN ACTIVITY-BASED TRAVEL MODEL**

### **Daehyun You (*corresponding author*)**

Arizona State University, School of Sustainable Engineering and the Built Environment  
Room ECG252, Tempe, AZ 85287-5306. Tel: (480) 965-3589; Fax: (480) 965-0557  
Email: [dae.you@asu.edu](mailto:dae.you@asu.edu)

### **Venu M. Garikapati**

Arizona State University, School of Sustainable Engineering and the Built Environment  
Room ECG252, Tempe, AZ 85287-5306. Tel: (480) 965-3589; Fax: (480) 965-0557  
Email: [venu.garikapati@asu.edu](mailto:venu.garikapati@asu.edu)

### **Ram M. Pendyala**

Arizona State University, School of Sustainable Engineering and the Built Environment  
Room ECG252, Tempe, AZ 85287-5306. Tel: (480) 727-9164; Fax: (480) 965-0557  
Email: [ram.pendyala@asu.edu](mailto:ram.pendyala@asu.edu)

### **Chandra R. Bhat**

The University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172  
Tel: (512) 471-4535, Fax: (512) 475-8744; Email: [bhat@mail.utexas.edu](mailto:bhat@mail.utexas.edu)

### **Subodh Dubey**

The University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712-1172  
Tel: (512) 471-4535, Fax: (512) 475-8744; Email: [subbits@gmail.com](mailto:subbits@gmail.com)

### **Kyunghwi Jeon**

Maricopa Association of Governments  
302 N First Avenue, Suite 300, Phoenix, AZ 85003. Tel: (602) 452-5079; Fax: (602) 254-6490  
Email: [kjeon@azmag.gov](mailto:kjeon@azmag.gov)

### **Vladimir Livshits**

Maricopa Association of Governments  
302 N First Avenue, Suite 300, Phoenix, AZ 85003. Tel: (602) 452-5079; Fax: (602) 254-6490  
Email: [vlivshits@azmag.gov](mailto:vlivshits@azmag.gov)

March 2014

## **ABSTRACT**

This paper describes the development of a vehicle fleet composition and utilization model system that may be incorporated into a larger activity-based travel demand model. It is of interest and importance to model household vehicle fleet composition and utilization behavior as the energy and environmental impacts of personal travel are not only dependent on the number of vehicles, but also on the mix of vehicles that a household owns and the extent to which different vehicles are utilized. A vehicle composition (fleet mix) and utilization model system has been developed for integration in the activity based travel demand model that is being developed for the Greater Phoenix metropolitan area in Arizona. At the heart of the vehicle fleet mix model system is a multiple discrete continuous extreme value (MDCEV) model capable of simulating vehicle ownership and use patterns of households. Vehicle choices are defined by a combination of vehicle body type and age category and the model system is capable of predicting household vehicle composition and utilization patterns at the household level. The paper describes the model system and presents results of a validation and policy sensitivity analysis exercise demonstrating the efficacy of the model.

**Keywords:** vehicle fleet composition, vehicle utilization, multiple discrete continuous extreme value (MDCEV) model, vehicle count modeling, travel demand forecasting

## INTRODUCTION

This paper is motivated by the recognition that the energy and environmental impacts of person travel are not only dependent on the number of vehicles, but on the mix of vehicle types that households own, and the extent to which households utilize vehicles of different types. Despite the widespread recognition of the importance of vehicle fleet composition and utilization behavior in transportation planning and modeling, such models have not yet found their way into operational activity-based travel demand model systems that are being implemented in metropolitan areas around the world. This paper is intended to fill this gap by presenting a comprehensive household vehicle fleet composition and utilization model system that may be integrated within a larger activity-based travel demand model. There has been considerable progress in the modeling of vehicle fleet composition and utilization in the recent past. This progress has been motivated by the fact that transportation accounted for 28% of greenhouse gas emissions (1) and 71% of all petroleum consumption in the United States in 2011 (2). In an effort to encourage households to acquire and use alternative and clean fuel vehicles, the government offers rebates and tax incentives tied to the purchase of such clean vehicles. The impacts of such policy actions on vehicle fleet composition, and the energy and emissions footprint of personal travel, can only be determined if travel model systems incorporate the ability to forecast vehicle fleet composition and utilization under a wide variety of socio-economic, built environment, and policy scenarios.

The emergence of the multiple discrete continuous extreme value (MDCEV) modeling methodology has provided a practical method for modeling vehicle fleet composition and utilization. A number of studies have employed this methodology to model household vehicle fleet mix (3-4). Vyas, et al (5) further extended the previous work in this domain to model not only vehicle fleet composition and utilization, but also the assignment of the primary driver of the household to each vehicle. Paleti, et al (6) describe a detailed vehicle fleet composition and evolution model system that may be applied to simulate household vehicle fleet mix over time. Several other studies have also attempted to model vehicle transactions behavior, albeit without explicit consideration of vehicle type choice behavior which is critical to forecasting vehicle fleet mix (7-8).

At the heart of the model system described in this paper is the MDCEV model, which is a methodological approach capable of modeling behavioral phenomena characterized by multiple discrete choices and a joint continuous choice dimension. In the vehicle fleet composition and utilization modeling context, households may choose to own multiple vehicles of different types (or zero vehicles or just one vehicle) and utilize (drive) each vehicle a different amount. These choice dimensions are modeled simultaneously within the MDCEV modeling framework. The next section presents the entire vehicle fleet composition and utilization framework that the study team has developed.

In addition to presenting the modeling framework, the paper presents model estimation results and a description of the data used for model development. The paper also includes results from a set of model validation and policy simulation exercises which demonstrate the operational feasibility and efficacy of the model system. Concluding thoughts and directions for further research are offered in the final section of the paper.

## VEHICLE FLEET COMPOSITION MODELING FRAMEWORK

The vehicle fleet mix modeling framework developed in this study is capable of simulating household vehicle fleet composition and utilization patterns while recognizing the multiple

discrete and continuous nature of the choice dimensions embedded in such a modeling effort. The modeling framework is depicted in Figure 1.

The first element of the model system is a household mileage prediction model. The MDCEV model requires a budget (a measure of the total continuous dimension available to the household) estimate that is then allocated to the various discrete alternatives chosen by the household. In the context of the vehicle fleet composition and utilization model, the budget is the total mileage driven or traveled by the household over the course of a year. This annual motorized mileage budget of the household is estimated using a linear regression model with a power transformation (equation 1) that avoids the sticky situation of negative mileage predictions that a standard linear regression model may provide.

$$(\text{Motorized Mileage})^{0.3} = \sum_{i=1}^k \beta_i x_i + \varepsilon \quad (1)$$

The value of 0.3 is obtained through a process of trial-and-error in an effort to identify the transformation that yields a distribution most amenable to being modeled using a linear regression. The mileage prediction model is followed by a MDCEV model that allocates the predicted budget differentially across the fleet mix chosen by a household. The application of the MDCEV model in forecasting mode entails the evaluation of multi-dimensional integrals of probability density functions. The computational intensity of these calculations had hindered the application of the MDCEV model in practice. However, the recent work of Pinjari and Bhat (9), who developed computationally efficient procedures for evaluating the multi-dimensional integrals using scrambled Halton sequences (10), has made it possible to apply the MDCEV model in practice.

The process of applying the MDCEV model involves running the simulation procedure repeatedly; each run offers a slightly different prediction of the vehicle fleet composition and utilization. An average may be computed over many runs of the simulation procedure, but there is no assurance that this approach is going to provide valid projections of vehicle fleet composition and utilization. From a practical standpoint, it is not possible to determine the number of runs of the simulation procedure that is sufficient to provide a robust and accurate forecast of vehicle fleet composition and utilization. How many runs of the simulation procedure must be executed before it can be stopped? This question presents a practical implementation challenge that remains the subject of ongoing research and study.

In order to overcome this challenge, the framework incorporates a simple multinomial logit (MNL) model of the number of vehicle body types owned by each household to provide marginal control totals to the MDCEV model. While the actual distribution of households by number of vehicle types is known in the base year (from survey data), the distribution is predicted using the MNL model for any future year. The MDCEV forecasting procedure is then run repeatedly until the MDCEV model predictions (of vehicle fleet composition and utilization) fall within an acceptable tolerance of the overall vehicle body type distribution predicted by the MNL model.

One of the limitations of the MDCEV model is that it does not provide an explicit count of vehicles that a household may own within a certain discrete alternative. For purposes of MDCEV model estimation and implementation, information for all vehicles that fall into the same discrete alternative is aggregated to form a single entity. Once the MDCEV model provides estimates of consumption of different vehicle types, separate count models can be run

for all consumed vehicle types to identify the number of vehicles within each consumed discrete alternative. In order to obtain the count of vehicles in each consumed category for each household, the framework includes a series of ordered probit count models. The question remains as to how the mileage allocated to a certain vehicle type category should be distributed across multiple vehicles in the same discrete category. At this time, once the ordered probit count model predicts the number of vehicles in each discrete category, the total mileage allocated to the vehicle type is equally distributed (shared) among all vehicles within the category.

The vehicle fleet composition and utilization model framework described in this section provides a practical approach for predicting vehicle fleet mix at the household level, including the number of vehicles by type owned by the household and the annual miles that each vehicle in the fleet is driven.

## **DATA DESCRIPTION**

This section presents a brief overview of the data used in the development of the vehicle fleet composition and utilization model system for the Greater Phoenix metropolitan area. Data used for model development is derived from the Greater Phoenix metropolitan area sample in the 2008-2009 National Household Travel Survey (NHTS). The NHTS is conducted periodically by the US Department of Transportation to collect data on socio-economic, demographic and personal travel characteristics of the nation. The NHTS also collects detailed data about all vehicles owned by households in the survey sample. For purposes of this effort, vehicles are classified into 13 categories, defined by a cross between four vehicle body types and three age categories plus the motorbike category. The four body types are car, van, SUV, and pick-up truck. The three age categories are 0-5, 6-11, and more than 11 years old. These age categories were chosen based on an analysis of the average age distribution of the vehicle fleet in the sample and with consideration that many vehicle manufacturers offer five-year power train warranty periods (11). The motorbike category serves as a thirteenth alternative. In order to accommodate zero-vehicle households, a non-motorized vehicle alternative was introduced. The non-motorized vehicle alternative is one that every household must consume and is treated as an outside good (an alternative consumed by all behavioral units) in the MDCEV model estimation effort. The non-motorized vehicle serves as the fourteenth alternative in the MDCEV model. Non-motorized vehicle mileage is computed both based on reported bicycle and walk trips and using the formula:  $0.5 \text{ mile per person per day} \times 365 \text{ days/year} \times \text{number of persons in household}$ . The larger of the two values is used as the estimate of non-motorized mileage for the household. This formulation has been used successfully in past studies (5). In general, MDCEV model estimates have been found to be robust and unaffected by the mileage allocation to this outside good (5). The MDCEV model has a total of 14 alternatives (12 vehicle alternatives + motorbike + non-motorized vehicle).

In the interest of brevity, a tabulation of socio-economic and demographic characteristics of the survey sample is not furnished in this paper. After extensive cleaning and analysis of the data, the estimation sample included 4,262 households. In general, the survey sample depicts socio-economic and demographic characteristics consistent with expectations. Table 1 provides a description of the vehicle fleet in the sample.

The survey sample data set was augmented with a number of built environment and accessibility variables provided by Maricopa Association of Governments (MAG). Using network skim data, accessibility measures were calculated for all traffic analysis zones (TAZs) including the amount of employment (of different categories) accessible within different travel

time bands of each TAZ. These regional employment accessibility measures were appended to the household records depending on their residential location.

### **ESTIMATION OF VEHICLE FLEET MIX MODEL SYSTEM COMPONENTS**

This section presents the various model components of the vehicle fleet composition model system and the model estimation results. Within the scope of this paper, it is impossible to provide detailed tabulations of estimation results for all model components. As such, detailed tabulations are provided only for the MDCEV model that is the heart of the vehicle fleet composition model system. Other model components are described in brief with mention of the key highlights in the model estimation results.

#### **Mileage Prediction Model**

The mileage prediction model is a power transformed linear regression model as depicted in equation (1). Model estimation results show that the mileage is influenced by socio-economic and demographic variables, as well as built environment and accessibility variables. The number of drivers in the household, number of adults in the household, and living in a rural area contribute positively to annual household mileage. Living in a rural area is likely to entail greater mileage due to the need to travel farther distances to access destination opportunities. Higher income households report greater mileage than lower income households. Households residing in TAZs with high levels of accessibility (i.e., in the top quartile of TAZs sorted by the amount of employment accessible within 10 minutes by auto) show a propensity for lower annual mileage. The model has a  $R^2$  value of 0.4, which is quite reasonable for a disaggregate household level regression model of annual mileage.

#### **MDCEV Model of Vehicle Fleet Composition and Utilization**

The MDCEV model is ideally suited to the modeling of vehicle fleet composition and utilization due to the multiple discrete and continuous nature of the behavioral phenomenon under study. As there is a fairly large and growing body of literature on the development and estimation of MDCEV models (9, 12), a detailed description of the methodology is not furnished in this paper in the interest of brevity. In a multiple discrete choice context, the MDCEV model formulation reduces the dimensionality of the problem (relative to a single discrete choice model such as the MNL model) through a parsimonious and efficient representation of the elemental alternatives in the choice set. In addition, the MDCEV model incorporates diminishing marginal returns (or satiation effects) in the consumption of an alternative, and integrates the modeling of a continuous choice dimension within the multiple discrete choice framework.

As the current empirical context involves the possibility that some households will choose to own absolutely no vehicles, the model specification must be formulated to accommodate this choice possibility. This is done by introducing the non-motorized vehicle as an outside good, an alternative that is chosen by all households in the sample. The functional form of the utility expression in the MDCEV model derived by Bhat (12) for a case with the presence of an outside good is as follows:

$$U(x) = \frac{1}{\alpha_{out}} \psi_{out} x_{out}^{\alpha_{out}} + \sum_{k=2}^K \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left( \frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \quad (2)$$

where  $\psi_{out} = \exp(\varepsilon_{out})$ , and  $\psi_k = \exp(\beta'z_k + \varepsilon_k)$ . The baseline marginal utility,  $\psi_k = \exp(\beta'z_k + \varepsilon_k)$  is to introduce the impact of observed and unobserved alternative attributes (12).  $z_k$  is a set of attributes characterizing the alternative  $k$  and  $\varepsilon_k$  captures unobserved characteristics that impact the baseline utility for good  $k$ .  $U(x)$  is a quasi-concave and continuously differentiable function with respect to consumption quantity vector  $x$  ( $x_k \geq 0 \forall k$ ).  $\psi_k$  represents the marginal utility for alternative  $k$  at the point of zero consumption.  $\alpha_k$  is the satiation parameter which governs variation in marginal utility with increasing consumption for good  $k$ . The translation parameter  $\gamma_k$  also governs the level of satiation, but more importantly enables corner solutions (i.e., zero consumption of some goods). Translation parameter  $\gamma_{out}$  is not estimated for the outside good, as it is always consumed and hence cannot have a corner solution. The  $\gamma$ -profile MDCEV model is estimated in which the satiation parameters  $\alpha_k$  of all alternatives are set to zero, and a separate translation parameter  $\gamma_k$  is estimated for each of the alternatives except the outside good. Under these restrictions, the model estimated in this study adopts the following utility formulation:

$$U(x) = \psi_{out} \ln(x_{out}) + \sum_{k=2}^K \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) \quad (3)$$

Model estimation results for the MDCEV model of vehicle fleet composition and utilization are furnished in Table 2. The outside good, non-motorized vehicle, serves as the base alternative in the model estimation. In general, model estimation results are behaviorally intuitive and consistent with expectations suggesting that the MDCEV model is a reasonable approach to modeling vehicle fleet composition and mileage. High income households are more likely to own newer vehicles and prefer cars and pick-up trucks over other types of vehicles. Households with children have a greater proclivity towards owning vans, which is consistent with the notion that vans are convenient for transporting families. Larger households tend to prefer vans, while households in rural areas exhibit a greater tendency to own trucks (relative to non-rural households). Households residing in single family dwelling units are more likely to own motorbikes; it is likely that these households own a motorbike as a recreational (hobby) vehicle as opposed to a regular means of transportation.

Among TAZ characteristics, it is found that households residing in TAZs with high proportion of households in the lowest income quintile were less likely to own new cars and SUVs. It is possible that spatial and social dependency effects play a role in vehicle ownership and this finding is consistent with such a notion (13). An examination of the influence of accessibility variables indicates that households in higher accessibility zones are less likely to own newer vehicles. As households in these high accessibility locations do not have to drive long distances, and can presumably use alternative modes of transport to satisfy their mobility needs, they can keep older vehicles in their fleet mix. Households in lower density areas, on the other hand, tend to own newer vehicles and have larger vehicles such as SUVs in their fleet mix.

A review of the baseline constants provides an indication of the inherent preferences for various vehicles, and the marginal utility at zero consumption for the different alternatives.

Among all alternatives considered, cars have the highest baseline utility and hence the greatest baseline preference. Among cars, there is a greater baseline preference for newer cars than for older cars, which is consistent with the actual patterns of vehicle ownership in the survey data set. Old vans have the smallest baseline utility suggesting that households generally prefer not to own such vehicles and derive the least utility from such vehicles (all other things being equal). In general baseline utility decreases with age of the vehicle, although this trend is not seen consistently for SUVs and pick-up trucks. In the case of SUVs, there is lower baseline utility for mid-aged vehicles suggesting that households tend to acquire newer SUVs and then hold on to their SUVs for a long time. In the case of pick-up trucks, there is a lower baseline utility for new pick-up trucks suggesting that households tend to acquire and hold older pick-up trucks and do not necessarily see the need to own newer pick-up trucks.

Translation parameters represent the diminishing marginal utility with increasing consumption of various alternatives and the extent to which households are inclined to drive various vehicle types. A higher translation parameter suggests that households are less satiated with the consumption of that alternative and are likely to drive that vehicle type more. In general, the translation parameters show a pattern where newer vehicles have larger translation parameter values than older vehicles. This is consistent with expectations that newer vehicles are generally driven more miles than older vehicles. New vans have the highest translation parameter, consistent with the notion that these vehicles serve as the multipurpose family vehicle driven for all purposes and to meet chauffeuring needs. Motorbikes have the lowest translation parameter, implying that these are a specialized vehicle type driven more sparingly than regular four-wheeled vehicles.

### **MNL Model of Number of Vehicle Types**

The proposed modeling framework incorporates a MNL model of number of vehicle types (owned by households) that provides a marginal control distributions to the MDCEV model should replicate. When the MDCEV model predictions replicate the marginal control distribution predicted by the MNL model (within a certain tolerance level), then the MDCEV forecasting procedure can be terminated and the predictions offered by the MDCEV model at that simulation run can be accepted as the forecast.

The MNL model includes five alternatives, corresponding to the four vehicle body types included in the model framework. Households may own zero, one, two, three, or four vehicle types. The forecast from the MDCEV model is aggregated and compared against the distribution predicted by the MNL model to determine if the MDCEV forecasting procedure can be terminated. The MNL model estimation results are not furnished in tabular form in the interest of brevity. It is found that households in high population density TAZs display a lower preference for owning multiple (different) vehicle body types. They tend to own fewer body types, reflecting a more homogeneous vehicle fleet composition. High income households, and households with larger number of workers, number of adults, or presence of children, tend to own a multiplicity of vehicle types, reflecting a more heterogeneous vehicle fleet composition. Households in rural areas tend to own two vehicle body types; this is consistent with the notion that such households may own a pick-up truck for utility purposes and a second vehicle type for personal travel purposes.



### **Ordered Probit Models of Vehicle Count by Type**

The MDCEV model of vehicle fleet composition and utilization is capable of providing the extent to which households own and utilize each vehicle type alternative (defined as a combination of vehicle body type and vintage). The MDCEV model does not, however, predict the actual count of vehicles within each body-type/age category. For example, if a household owns two cars in age category 0-5 years old, the MDCEV model predicts and attributes all of the mileage consumption of the household to car 0-5 years old category without providing any indication of how many such vehicles are owned by the household. In order to complete the vehicle composition and utilization profile of a household, vehicle count models are needed. In the proposed framework, ordered probit models of vehicle count by type are estimated and implemented to serve this purpose as these models fit the data best among various types of count models tested. The ordered probit model of a specific vehicle body type is applied only when a household is predicted by the MDCEV model to have a non-zero consumption in that specific vehicle type category.

Four different ordered probit models were estimated in this modeling framework. Although it is feasible to have 12 (or 13, including motorbike) ordered probit models, one for each vehicle body/age alternative, it was felt that the inclusion of so many ordered probit models would make the model system difficult to calibrate and validate as there would be many parameters across the model system. As a reasonable compromise, four ordered probit models are included in the model system, one for each vehicle body type. Thus, the “car” ordered probit model is applied to determine the count of cars in any age category of car for which the MDCEV model predicts a non-zero consumption. Age indicators enter the utility expression for each vehicle type to reflect the differential effects of vehicle age on ownership patterns. In the interest of brevity, the four ordered probit vehicle count models are not presented in tabular form in this paper.

It is found that all four ordered probit models include all three age indicators as significant explanatory variables in predicting count of vehicles of a certain type. The annual mileage also appears as an explanatory variable in the count model; presumably, if the mileage is very large, then it is likely that the household owns multiple vehicles of the said type as it is unlikely that a single vehicle will be allocated a large mileage budget. These findings are consistent with expectations and with those found in the MNL model. These variables positively contribute to vehicle count across all four body types.

### **SAMPLE REPLICATION AND MODEL SENSITIVITY TEST**

The entire model system was calibrated and tested by undertaking a sample replication and model sensitivity exercise. First, the model system was applied to predict vehicle ownership, composition, and utilization patterns for each household in the entire estimation sample. This process does not constitute a true validation exercise. In a true validation exercise, the survey sample would have been split into an estimation sample and a validation sample (hold-out sample of 20 or 30 percent of the entire survey sample). However, the number of components included in the model system, and the use of 14 alternatives in the MDCEV model system, necessitated the use of the entire survey sample of 4,262 households in the model estimation process. The estimated models were therefore applied to the estimation sample and the predicted patterns of vehicle ownership were compared against the actual observed patterns to determine the ability of the model system to replicate patterns of behavior in the estimation sample. The sample replication process was used to determine any discrepancies in the model’s predictive

ability and make fine adjustments to model constants and parameters so that the model system was able to closely replicate observed vehicle ownership and utilization patterns.

The MDCEV forecasting procedure was implemented using the Gauss codes that have been made available by Pinjari and Bhat (9). In general, it was found that the uncalibrated model was able to predict the annual mileage allocation across vehicle types quite well. The model was also able to predict the percent of households owning different vehicle types reasonably well, although there was a systematic tendency on the part of the model to slightly under-predict the percent of households owning different vehicle types (virtually across all vehicle type alternatives). In order to improve the match further, a few baseline constants and translation parameters were fine-tuned through a cautious calibration exercise so that the model predictions more closely matched observed patterns in the estimation data set. The top half of Table 3 provides results of the calibration effort.

It can be seen that the percent differences in vehicle ownership by type and average annual mileage by vehicle class are quite small, suggesting that the calibrated MDCEV model is able to closely replicate patterns of vehicle ownership in the survey sample. Extensive comparisons were also performed within specific market segments; it was found that the model replicated vehicle ownership patterns closely within all key market segments (tabulations suppressed).

Following the calibration of the MDCEV model, the study team examined the ability of the MNL model to replicate the overall distribution of households by number of vehicle types, both in the aggregate and for a number of key market segments (households in urban areas, households in rural areas, households in different income quintiles, households with and without children, for example). In general, it was found that the model performed extremely well in replicating overall vehicle type count ownership patterns with virtually no need for any calibration adjustments. The comparison of observed and predicted patterns of vehicle type count is shown at the household level in Table 4. In general, it is found that the MNL model produces a distribution of households by number of vehicle types that closely mirrors that seen in the survey sample data set. A similar exercise was performed for the ordered probit models of vehicle counts within each vehicle body type. Once again, it was found that the ordered probit models were able to replicate observed count patterns very closely, with virtually no need for calibration of constants or thresholds. Results of this replication effort are shown in Table 4.

Once the model was calibrated and shown to perform well in sample replication, it was considered prudent to subject the model to a sensitivity test to examine the ability of the model system to respond in a meaningful way when input conditions are changed. The entire model system was applied to the survey sample of 4,262 households to establish baseline conditions (of mileage and vehicle ownership by type). Then, five different scenarios were created. The regional employment accessible by auto is incrementally increased by 10%, 20%, 30%, 50%, and 100% for all TAZs in the region. This essentially means that total employment in the region is being increased substantially and that the higher levels of employment are realized in a denser pattern of development (than that which prevails in the baseline).

The results of this sensitivity test are presented in the bottom half of Table 3 to explain the trends in vehicle ownership with increasing accessibility. The table first shows the number of households owning each vehicle type in a particular scenario, and the value in parenthesis is the percent change in the number of households owning a vehicle type relative to the baseline (first row). The share of households owning each vehicle type can be calculated by dividing the values in these cells by the sample size (4,262 households). In general, the model provides

predictions of change in vehicle ownership and utilization patterns consistent with expectations. It is found that, as accessibility increases, the percent of households owning cars (smaller vehicles) increases while the percent of households owning larger vehicles (SUVs) steadily decreases. The percent of households owning vans drops as well, but not as much as SUVs, suggesting that households are more inclined to hold these multipurpose family vehicles in the fleet. The percent of households owning pick-up trucks remains largely unchanged. At very high levels of accessibility increase, a more noticeable drop in van ownership is discernible. These patterns of change are consistent with the notion that, as accessibility increases, households need to drive smaller distances to access employment and destination opportunities. As larger vehicles are generally preferred for longer trips (for comfort reasons) (14), they are no longer the preferred vehicle type in the event of increased accessibility. A rather interesting finding is that the percent of households owning motorbike increases substantially; these larger percent changes should be viewed with caution – as the baseline conditions involved a lower motorbike ownership level to begin with, any change will be amplified when percent changes are computed.

Second, the table presents the trends in mileage consumption with increasing accessibility. The results are organized in a similar fashion as the first part, where average annual mileage consumption for different vehicle types is provided for each scenario and percent change in consumption with respect to the baseline is given in parenthesis. As expected, average mileage values decrease for all car body types as accessibility increases, albeit to varying degrees. Car mileage gradually decreases with increasing accessibility, despite the trend of increasing market share. This finding is reflective of a combination effect; as accessibility increases, there is a move towards greater car ownership (thus increasing share), but an associated decrease in mileage (because the cars are driven shorter distances when accessibility increases). The mileage values for vans, SUVs, and pick-up trucks also show a decrease from the baseline scenario, although the trends are not as consistent as in the case of cars. The mileage is computed by dividing the total mileage attributed to a vehicle type by the number of vehicles in the fleet (in any given scenario). As the number of vehicles in the fleet drops, the smaller denominator in this calculation may contribute to a more modest decrease or even a modest increase in per vehicle type mileage values. In the case of vans and SUVs, for example, it is found that there is a decrease in market share that is more marked than the decrease in mileage values. These households (which retain the larger vehicles) presumably engage in longer distance travel (recreational travel, for example); they therefore hold these vehicles in the fleet and drive them about the same distance despite the increase in local accessibility. For non-motorized vehicles, the changes in mileage are comparable to the baseline because all households in the dataset consume this alternative in all scenarios (as it is the outside good). The rather large percent changes maybe a manifestation of the smaller baseline values that form the basis of percent change computations. The increase in non-motorized mileage may be readily explained. An increase in density and accessibility is likely to be associated with greater levels of bicycling and walking as distances to destination opportunities become smaller.

## CONCLUSIONS

This paper describes a comprehensive vehicle fleet composition and utilization model system that may be integrated in an activity-based travel demand model system. Concerns regarding energy sustainability and greenhouse gas emissions motivate the development of such a model system so that travel demand models are able to offer rich and detailed information to inform energy and emissions models. Armed with information about the exact mix of vehicles on the

roadways and the extent to which different vehicles are being driven, energy and environmental scientists will be able to better estimate energy consumption patterns and emissions inventories (in space and time) under alternative policy and socio-economic scenarios.

The model system developed in this study includes three primary components and one support component. The model system includes a power-transformed linear regression model to estimate the annual mileage of households. Next, the model system includes a MDCEV model to simulate the vehicle fleet mix and usage patterns at the household level. The third component of the model system is a series of ordered probit vehicle count models so that the precise count of vehicles within each body/age vehicle type alternative may be determined. Thus, the three primary components are capable of providing a complete picture of vehicle fleet composition, count, and mileage at the household level. A support component is introduced to help inform the simulation-based MDCEV forecasting application. A MNL model of number of vehicle body types owned by a household is estimated and applied to predict the number of households that own different numbers of vehicle types. This distribution is used as the basis to determine when the simulation-based MDCEV forecasting procedure has furnished an acceptable forecast of vehicle fleet composition and can be terminated.

The entire model system is applied (in sample replication mode) to a survey sample of 4,262 households for the Greater Phoenix metropolitan area. The models are estimated and calibrated using this survey sample of households. The model system is found to perform very well in replicating household vehicle ownership, composition, and utilization patterns. Then, the model is applied to scenarios where regional accessibility measures are increased across the entire region. The model system is found to respond in a reasonable manner in the event of a change in regional accessibility measures. As regional accessibility increases, larger vehicles are consumed less (both in ownership and utilization), and smaller vehicles (cars) are consumed more (primarily in ownership). As the higher levels of accessibility are likely to be associated with shorter distance trips to access destinations, the mileages of almost all the motorized modes is observed to decrease. As expected, regional accessibility increases are associated with an increase in bicycle and walk mileage. Overall, the model system appears to be suitable for implementation and use in an activity-based travel demand model that purports to simulate activity-travel choices at the level of the individual household and traveler.

The model system may be improved in a number of ways. The inclusion of fuel price, vehicle operating costs, and other vehicle attributes would enhance the policy sensitivity of the model system. The model system can be further enhanced to include a choice model capable of simulating the precise make/model/year of the vehicle(s) owned by households. Such disaggregate information about the vehicle fleet mix would further aid in enhancing the resolution of energy and environmental studies. The choice set currently captures the vehicle body type and age dimensions, but does not capture the fuel type dimension. The survey sample data set used in this study did not include enough of a variation in fuel type to capture that dimension. As the market share of alternative and clean fuel vehicles continues to increase, survey samples should include larger numbers of such households making it possible to include that dimension. Given the policy and technology implications of alternative and clean fuel vehicles, it is imperative that this dimension be captured in future updates of the model system.

**REFERENCES**

1. EPA, *Sources of Greenhouse Gas Emissions*. July 30, 2013. <http://www.epa.gov/climatechange/ghgemissions/sources/transportation.html>. Accessed July 31, 2013.
2. EIA, *Oil: Crude and Petroleum Products Explained*. Apr. 23, 2012. [http://www.eia.gov/energyexplained/index.cfm?page=oil\\_home#tab2](http://www.eia.gov/energyexplained/index.cfm?page=oil_home#tab2). Accessed July 31, 2013.
3. Bhat, C. R., S. Sen, and N. Eluru. The Impact of Demographics, Built Environment Attributes, Vehicle Characteristics, and Gasoline Prices on Household Vehicle Holdings and Use. *Transportation Research Part B*, Vol. 43, No. 1, 2009, pp. 1-18.
4. Eluru, N., C. R. Bhat, R. M. Pendyala, and K. C. Konduri. A Joint Flexible Econometric Model System of Household Residential Location and Vehicle Fleet Composition/Usage Choices. *Transportation*, Vol. 37, No. 4, 2010, pp. 603-626.
5. Vyas, G., R. Paleti, C. R. Bhat, K. G. Goulias, R. M. Pendyala, H. Hu, T. J. Adler, and A. Bahreinian. Joint Vehicle Holdings, by Type and Vintage, and Primary Driver Assignment Model with Application for California. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2302, 2012, pp. 74-83.
6. Paleti, R., N. Eluru, C. R. Bhat, R. M. Pendyala, T. J. Adler, and K. G. Goulias. Design of Comprehensive Microsimulator of Household Vehicle Fleet Composition, Utilization, and Evolution. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2254, 2011, pp. 44-57.
7. Mohammadian, A., and E. J. Miller. Dynamic Modeling of Household Automobile Transactions. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1831, 2003, pp. 98-105.
8. Yamamoto, T., J. Madre, and R. Kitamura. An Analysis of the Effects of French Vehicle Inspection Program and Grant for Scrapage on Household Vehicle Transaction. *Transportation Research Part B*, Vol. 38, No. 10, 2004, pp. 905-926.
9. Pinjari, A. R., and C. R. Bhat. A Multiple Discrete-continuous Nested Extreme Value (MDCNEV) Model: Formulation and Application to Non-worker Activity Time-use and Timing Behavior on Weekdays. *Transportation Research Part B*, Vol. 44, No. 4, 2010, pp. 562-583.
10. Bhat, C. R. Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. *Transportation Research Part B*, Vol. 37, No. 9, 2003, pp. 837-55.
11. My Car Stats. *Manufacturer Car Warranty Expiration Chart*. Oct. 2010. [http://www.mycarstats.com/content/car\\_warrantylimits.asp](http://www.mycarstats.com/content/car_warrantylimits.asp). Accessed July 31, 2013.
12. Bhat, C. R. The Multiple Discrete-continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions. *Transportation Research Part B*, Vol. 42, No. 3, 2008, pp. 274-303.
13. Paleti, R., C.R. Bhat, and R.M. Pendyala. The Modeling of Household Vehicle Type Choice Accommodating Spatial Dependence Effects. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2343, 2013, pp. 86-94.
14. Konduri, K. C., X. Ye, B. Sana, and R. M. Pendyala. Joint Model of Vehicle Type Choice and Tour Length. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2255, 2011, pp. 28-37.

**List of Tables**

<b>Caption</b>	<b>Table Title</b>
TABLE 1	Vehicle Characteristics
TABLE 2	MDCEV Model of Vehicle Fleet Composition and Utilization
TABLE 3	Comparison of Observed and MDCEV Model Predicted Vehicle Fleet Composition and Mileage
TABLE 4	Comparison of Observed and Predicted Vehicle Type and Vehicle Counts

**List of Figures**

<b>Caption</b>	<b>Figure Title</b>
FIGURE 1	Vehicle fleet composition and utilization model system.

**TABLE 1 Vehicle Characteristics**

	<b>Vehicle Body Type</b>				
	<b>Car</b>	<b>Van</b>	<b>SUV</b>	<b>Pick-up</b>	<b>Motor Bike</b>
Average Age (years)	8.55	7.46	6.52	9.52	9.21
Average Annual Mileage	10204.4	11317.7	11296.6	10723.0	3838.9
Number of Vehicles	3,997	635	1,537	1,376	240
<b>Vehicle Type vs. Annual mileage</b>					
<i>Annual Mileage</i>					
0 - 4,999	27.5%	18.4%	21.1%	24.9%	71.3%
5,000 - 9,999	30.6%	31.3%	28.6%	29.4%	15.8%
10,000 - 14,999	21.4%	26.9%	26.2%	22.8%	7.9%
15,000 - 19,999	11.3%	13.9%	12.8%	12.5%	2.9%
≥ 20,000	9.1%	9.4%	11.3%	10.5%	2.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%
<b>Vehicle Type vs. Vehicle Age</b>					
<i>Vehicle Age</i>					
0 - 5 Years	42.0%	40.9%	52.4%	34.2%	43.3%
6 - 11 Years	35.2%	44.3%	35.3%	39.5%	35.4%
≥ 12 Years	22.8%	14.8%	12.2%	26.3%	21.3%
Total	100.0%	100.0%	99.9%	100.0%	100.0%

**TABLE 2 MDCEV Model of Vehicle Fleet Composition and Utilization**

<b>Baseline Utility</b>			
<b>Explanatory Variable</b>	<b>Coef (t-stat)</b>	<b>Explanatory Variable</b>	<b>Coef (t-stat)</b>
<b><i>Car (0 - 5 years)</i></b>		<b><i>Van (12 years or older)</i></b>	
High income household (\$75,000 - \$99,999)	0.16 (2.16)	Lowest income household (< \$25,000)	0.66 (2.57)
Three or more workers in the household	0.17 (1.38)	Number of household members	0.14 (1.98)
Number of children in the household	-0.19 (-5.91)	TAZ with 1st (high) density quartile	0.68 (3.10)
Proportion of hhlds in the lowest income quintile	-1.01 (-3.61)	<b><i>SUV (0 - 5 years)</i></b>	
% of employment within 10 mins of accessibility	-13.83 (-3.24)	Lowest income household (< \$25,000)	-1.01 (-5.59)
<b><i>Car (6 - 11 years)</i></b>		Two workers in the household	0.17 (1.92)
Low income household (\$25,000 - \$49,999)	0.13 (1.90)	Household has retired adult(s) and no children	-0.17 (-1.78)
Two workers in the household	-0.16 (-2.21)	Proportion of hhlds in the lowest income quintile	-1.48 (-3.85)
<b><i>Car (12 years or older)</i></b>		TAZ with 3rd (low) density quartile	0.27 (2.48)
Lowest income household (< \$25,000)	0.57 (5.84)	% of employment within 10 mins of accessibility	-17.83 (-2.82)
Household has retired adult(s) and no children	0.20 (2.59)	<b><i>SUV (6 - 11 years)</i></b>	
Proportion of hhlds in the lowest income quintile	0.57 (1.74)	Medium income household (\$50,000 - \$74,999)	0.26 (2.41)
<b><i>Van (0 - 5 years)</i></b>		Household size = 4 or more	0.33 (3.30)
Two workers in the household	-0.39 (-2.53)	Household in an owned single family housing unit	0.86 (4.37)
Number of children in the household	0.38 (8.23)	TAZ with highest accessibility quartile (30 mins)	-0.30 (-2.85)
TAZ with 1st (high) density quartile	-0.31 (-1.89)	<b><i>SUV (12 years or older)</i></b>	
% of employment within 30 mins of accessibility	-1.72 (-1.78)	High income household (\$75,000 - \$99,999)	0.36 (1.89)
<b><i>Van (6 - 11 years)</i></b>		Presence of children in the household	0.31 (1.99)
Low income household (\$25,000 - \$49,999)	0.27 (1.87)	Household in a single family housing unit	-0.74 (-2.38)
Number of children in the household	0.33 (7.15)	TAZ with 2nd (medium) density quartile	-0.36 (-2.28)
TAZ with 1st (high) density quartile	-0.26 (-1.87)	<b><i>Motor Bike</i></b>	
		Household size = 1	-0.57 (-2.11)
		Household resides in rural area	0.71 (4.41)
		Household in an owned single family housing unit	0.75 (2.36)



**TABLE 2 MDCEV Model of Vehicle Fleet Composition and Utilization (continued)**

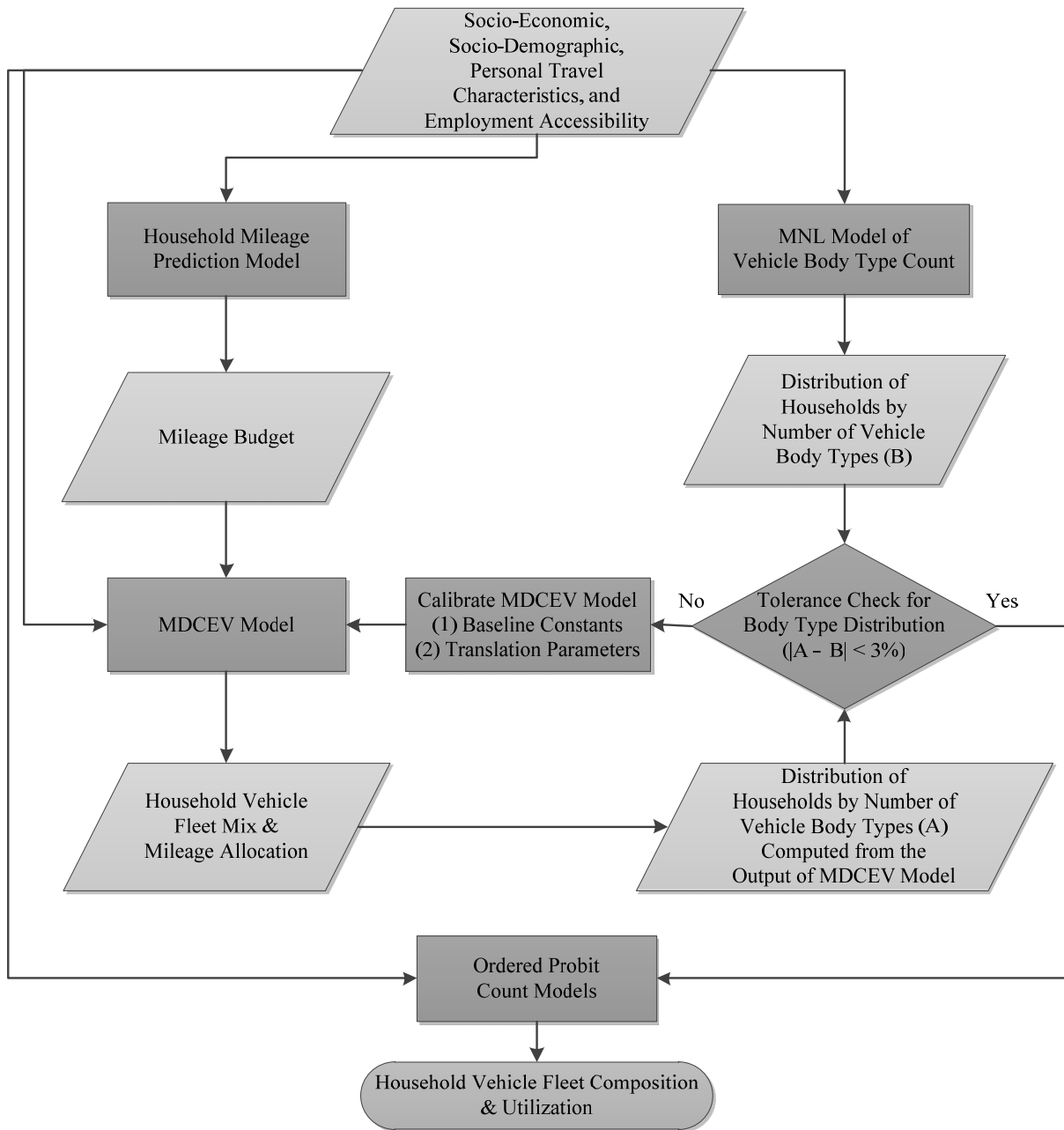
<b>Baseline Utility</b>			
<b>Explanatory Variable</b>	<b>Coef (t-stat)</b>	<b>Explanatory Variable</b>	<b>Coef (t-stat)</b>
<i>Pick-up (0 - 5 years)</i>		<i>Pick-up (12 years or older)</i>	
Highest income household (>= \$100,000)	0.24 (2.28)	Low income household (\$25,000 - \$49,999)	0.35 (2.86)
Household size = 1	-0.98 (-4.64)	Presence of children in the household	-0.25 (-1.95)
Household resides in rural area	0.24 (1.95)	Proportion of hhlds in the lowest income quintile	1.28 (2.95)
Proportion of single family housing units in the TAZ	0.74 (2.44)	TAZ with highest accessibility quartile (10 mins)	-0.34 (-2.60)
<i>Pick-up (6 - 11 years)</i>		<b>Goodness of fit</b>	
High income household (\$75,000 - \$99,999)	0.16 (1.44)	Log-likelihood of base model at convergence (df:25)	-77343.17
Household has retired adult(s) and no children	-0.35 (-3.41)	Log-likelihood of final model at convergence (df:75)	-77020.49
Household resides in rural area	0.15 (1.30)	Likelihood ratio	645.36
TAZ with highest accessibility quartile (10 mins)	-0.23 (-2.17)	$\chi^2_{50, .001}$	86.66
<b>Baseline Constants</b>		<b>Translation Parameters</b>	
<b>Vehicle Type</b>	<b>Coef (t-stat)</b>	<b>Vehicle Type</b>	<b>Coef (t-stat)</b>
Car 0-5 years old	-5.98 (-83.27)	Non-motorized vehicle (Outside Good)	0
Car 6-11 years old	-6.51 (-140.27)	Car 0-5 years old	23668 (10.07)
Car 12 years or older	-7.29 (-93.28)	Car 6-11 years old	18621 (10.37)
Van 0-5 years old	-8.13 (-61.11)	Car 12 years or older	12164 (9.46)
Van 6-11 years old	-8.43 (-82.80)	Van 0-5 years old	29431 (3.41)
Van 12 years or older	-10.04 (-41.51)	Van 6-11 years old	22248 (4.21)
SUV 0-5 years old	-6.65 (-64.00)	Van 12 years or older	12691 (3.22)
SUV 6-11 years old	-8.32 (-42.25)	SUV 0-5 years old	25172 (6.70)
SUV 12 years or older	-7.88 (-25.94)	SUV 6-11 years old	16717 (6.71)
Pick-up 0-5 years old	-8.29 (-30.87)	SUV 12 years or older	8397 (5.10)
Pick-up 6-11 years old	-7.35 (-95.28)	Pick-up 0-5 years old	20610 (5.69)
Pick-up 12 years or older	-8.06 (-70.37)	Pick-up 6-11 years old	14758 (6.92)
Motor Bike	-9.24 (-29.10)	Pick-up 12 years or older	9541.7 (6.70)
		Motor Bike	2223.2 (7.67)

**TABLE 3 Comparison of Observed and MDCEV Model Predicted Vehicle Fleet Composition and Mileage**

<b>Results of Validation/Replication Exercise</b>														
<b>Vehicle Type Distribution (percent of households owning vehicle type)</b>														
	<b>Non Motor</b>	<b>Car 0-5</b>	<b>Car 6-11</b>	<b>Car &gt;=12</b>	<b>Van 0-5</b>	<b>Van 6-11</b>	<b>Van &gt;=12</b>	<b>SUV 0-5</b>	<b>SUV 6-11</b>	<b>SUV &gt;=12</b>	<b>Pick-up 0-5</b>	<b>Pick-up 6-11</b>	<b>Pick-up &gt;=12</b>	<b>Motor Bike</b>
Observed	100.0%	34.8%	30.1%	18.9%	6.1%	6.4%	2.2%	17.6%	12.3%	4.2%	10.5%	12.2%	8.1%	4.6%
Predicted	100.0%	34.7%	29.4%	19.2%	4.6%	5.9%	2.2%	17.6%	13.2%	3.8%	11.3%	11.3%	7.7%	4.5%
Difference (%)	0.0%	0.2%	0.6%	-0.3%	1.4%	0.5%	0.0%	0.0%	-0.9%	0.5%	-0.7%	0.8%	0.4%	0.1%
<b>Average Annual Mileage (miles)</b>														
Observed	416	13218	10951	8863	12432	11338	9350	13531	11059	7707	13461	10870	8928	4653
Predicted	435	13337	11160	9069	12631	11354	9753	13693	11072	7978	13526	11437	8491	4901
Difference (miles)	-19	-119	-209	-206	-199	-16	-403	-162	-13	-271	-66	-566	437	-248
Difference (%)	-4.7%	-0.9%	-1.9%	-2.3%	-1.6%	-0.1%	-4.3%	-1.2%	-0.1%	-3.5%	-0.5%	-5.2%	4.9%	-5.3%
<b>Results of Model Sensitivity Test</b>														
<b>Change of Vehicle Ownership</b>														
<b>Scenarios for Sensitivity Analysis</b>	<b>Non Motor</b>	<b>Car</b>	<b>Van</b>	<b>SUV</b>	<b>Pick-up</b>	<b>Motor Bike</b>								
	<b>Count (Δ%)</b>	<b>Count (Δ%)</b>	<b>Count (Δ%)</b>	<b>Count (Δ%)</b>	<b>Count (Δ%)</b>	<b>Count (Δ%)</b>								
Baseline	4262 (0.0%)	3676 (0.0%)	566 (0.0%)	1521 (0.0%)	1300 (0.0%)	189 (0.0%)								
Accessibility (↑) 10%	4262 (0.0%)	3682 (0.2%)	568 (0.4%)	1508 (-0.9%)	1300 (0.0%)	190 (0.5%)								
Accessibility (↑) 20%	4262 (0.0%)	3701 (0.7%)	564 (-0.4%)	1500 (-1.4%)	1300 (0.0%)	191 (1.1%)								
Accessibility (↑) 30%	4262 (0.0%)	3711 (1.0%)	565 (-0.2%)	1493 (-1.8%)	1296 (-0.3%)	193 (2.1%)								
Accessibility (↑) 50%	4262 (0.0%)	3718 (1.1%)	561 (-0.9%)	1479 (-2.8%)	1303 (0.2%)	195 (3.2%)								
Accessibility (↑) 100%	4262 (0.0%)	3734 (1.6%)	554 (-2.1%)	1442 (-5.2%)	1308 (0.6%)	202 (6.9%)								
<b>Change of Vehicle Annual Mileage</b>														
	<b>Mileage (Δ%)</b>	<b>Mileage (Δ%)</b>	<b>Mileage (Δ%)</b>	<b>Mileage (Δ%)</b>	<b>Mileage (Δ%)</b>	<b>Mileage (Δ%)</b>								
Baseline	475 (0.0%)	11600 (0.0%)	11780 (0.0%)	11879 (0.0%)	11635 (0.0%)	4739 (0.0%)								
Accessibility (↑) 10%	478 (0.7%)	11609 (0.1%)	11751 (-0.3%)	11819 (-0.5%)	11575 (-0.5%)	4751 (0.3%)								
Accessibility (↑) 20%	481 (1.4%)	11578 (-0.2%)	11785 (0.0%)	11760 (-1.0%)	11560 (-0.6%)	4745 (0.1%)								
Accessibility (↑) 30%	483 (1.8%)	11569 (-0.3%)	11701 (-0.7%)	11759 (-1.0%)	11572 (-0.5%)	4755 (0.3%)								
Accessibility (↑) 50%	488 (2.8%)	11550 (-0.4%)	11695 (-0.7%)	11781 (-0.8%)	11492 (-1.2%)	4762 (0.5%)								
Accessibility (↑) 100%	500 (5.3%)	11525 (-0.6%)	11806 (0.2%)	11857 (-0.2%)	11482 (-1.3%)	4722 (-0.4%)								

**TABLE 4 Comparison of Observed and Predicted Vehicle Type and Vehicle Counts  
Vehicle Type Count in Household (Multinomial Logit Model)**

	<i>0 Vehicles</i>	<i>1 Type</i>	<i>2 Types</i>	<i>3 Types</i>	<i>4 Types</i>
Observed	5.1%	49.7%	40.1%	4.9%	0.1%
Predicted	4.5%	49.9%	40.5%	5.0%	0.1%
Difference (%)	0.6%	-0.2%	-0.4%	-0.1%	0.0%
<b>Vehicle Ownership Count within Vehicle Body Type (Ordered Probit Models)</b>					
<i>Car</i>	<i>1 Car</i>	<i>2 Cars</i>	<i>3 Cars +</i>		
Observed	72.6%	23.1%	4.3%		
Predicted	72.8%	23.2%	3.9%		
Difference (%)	-0.2%	-0.1%	0.4%		
<i>Van</i>	<i>1 Van</i>	<i>2 Vans +</i>			
Observed	95.0%	5.0%			
Predicted	95.4%	4.6%			
Difference (%)	-0.3%	0.3%			
<i>SUV</i>	<i>1 SUV</i>	<i>2 SUVs +</i>			
Observed	87.2%	12.8%			
Predicted	88.3%	11.7%			
Difference (%)	-1.1%	1.1%			
<i>Pick-up</i>	<i>1 Pick-up</i>	<i>2 Pick-ups +</i>			
Observed	89.0%	11.0%			
Predicted	87.9%	12.1%			
Difference (%)	1.1%	-1.1%			



**FIGURE 1 Vehicle fleet composition and utilization model system.**