Title
A Vehicle Usage Forecasting Model Based on Revealed and Stated Vehicle Type Choice and Utilization Data

Authors
Golob, Thomas F.
Bunch, David S.
Brownstone, David

Publication Date
1995-12-01

UCI-ITS-WP-95-23

A Vehicle Usage Forecasting Model Based on Revealed and Stated Vehicle Type Choice and Utilization Data

UCI-ITS-WP-95-23

Thomas F. Golob [1]
David S. Bunch [2]
David Brownstone [3]

[1] Institute of Transportation Studies, University of California, Irvine

[2] Graduate School of Management, University of California, Davis

[3] Department of Economics and Institute of Transportation Studies
University of California, Irvine

December 1995

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

1. OBJECTIVES AND RESEARCH CONTEXT

This research describes a new model of household vehicle usage behavior *by type of vehicle*. Forecasts of future vehicle emissions, including potential gains that might be attributed to introductions of alternative-fuel (clean-fuel) vehicles, critically depend upon the ability to forecast vehicle miles traveled by the fuel type, body style and size, and vintage of the vehicle.

Households acquire vehicles to satisfy both the transportation needs and the preferences of household members. Consequently, vehicle usage by type of vehicle can be considered to be a function of three categories of variables: (1) household characteristics, (2) principal driver characteristics, and (3) characteristics of the vehicle itself. Examples of *household characteristics* are income, residential location, number of vehicles, number of license-holders, number of workers, and number of household members by age group.

Usage of a specific vehicle depends heavily on which household license-holder typically drives the vehicle, i.e., the principal driver. Important *principal driver characteristics* include age, gender, and employment status. Workers, younger persons, and males are likely to drive more, as demonstrated in the models of Hensher (1985), Hensher, *et al.* (1992), Mannering (1983), Mannering and Winston (1985), and Train (1986).

Usage is also affected by *vehicle characteristics*, such as vehicle age (vintage), operating and capital costs, passenger and cargo capacity, and body style. Moreover, alternative-fuel vehicles are distinguished by vehicle attributes that potentially have an important influence on usage patterns; examples include limited range between refueling, coupled with limited fuel availability or the necessity to refuel or recharge the vehicle at home overnight. Differences between conventional-fuel and alternative-fuel vehicles in terms of fuel costs, cargo capacity, performance, and image are also expected to influence vehicle usage (van Wissen and Golob, 1992).

Applying a vehicle-type usage model in travel demand forecasts requires obtaining or developing forecasts of all of the model's exogenous variables. The first category of variables, *household characteristics*, can be readily forecast using Census data or household sociodemographic models used in regional planning. For example, the usage model developed here is part of a microsimulation forecasting system (Brownstone, *et al.*, 1994; Bunch, *et al.,* 1995) that is driven by a competing-risk hazard model of changing household demographics (Kazimi, 1995; Kazimi and Brownstone, 1995).

Forecasts of the second category of variables, characteristics of the *principal driver*, are problematic for multi-vehicle households, and also for single-vehicle households with more than one license-holder. For such households, modeling vehicle usage behavior involves allocating vehicles to license-holders in order to satisfy their activity needs (Golob, *et al.,* 1995). In addition, it involves distributing total travel among the vehicles and license-holders. While, in principle, we only need forecasts of household and vehicle characteristics to forecast vehicle usage for single-vehicle households with only one license-holder, we need exogenous forecasts of the characteristics of the principal driver of each vehicle in multi-vehicle and multi-driver households. To address this issue, the models described in this paper simultaneously incorporate allocation of license-holders to vehicles along with vehicle utilization.

Finally, exogenous forecasts of household vehicle holdings by *type of vehicle* are obtainable using vehicle type choice models, such as those developed by Lave and Train (1979), Manski and Sherman (1980), Hensher and Manefield (1982), Hocherman, et al. (1983), Berkovec (1985), Hensher and Le Plastrier (1985), Mannering and Winston (1985), Train (1986), McCarthy and Tay (1989), Hensher, *et al.* (1992). Such vehicle type choice models are based on vehicle holdings and transactions data (so called "revealed-preference" or RP models). Because consumers do not have actual experience with alternative-fuel vehicles that are likely to be available in 1998 and beyond, a vehicle type choice model based on stated preference (SP) data is required to forecast demand for these new vehicle types. One such model (Brownstone, *et al.*,

1995) is being coupled with the usage models described in this paper to forecast alternative-fuel vehicle usage for the State of California.

These models are similar to previous models of vehicle allocation and use in multi-vehicle households (Mannering, 1983; Hensher, 1985; Train, 1986; and Hensher, *et al.*, 1992) in that separate equations with correlated error terms are developed for each vehicle in the household. However, this research differs from previous efforts because there are additional equations that describe the most important characteristics of the principal driver of each vehicle. Although these characteristics cannot be readily forecast for use in a microsimulation system, they can be "solved out" of the problem; reduced-form equations are developed for forecasting purposes through a structural specification of vehicle allocation to drivers. Another unique feature is that the models use both RP and SP usage data simultaneously. In other words, the models are estimated with a mix of RP and SP observations, making the models sensitive to attributes associated with future alternative fuel vehicles.

The current version of the models takes the household's vehicle holdings (both the number of vehicles and the vehicle types) as given. This model structure is theoretically unappealing (as described in Golob, *et al.*, 1995), because a household's anticipated travel behavior is likely to influence its vehicle choices If the error terms of the vehicle choice model and the vehicle usage model are correlated, the parameter estimates will be biased. One approach is to apply a linear correction term involving a transformation of predicted vehicle choice probabilities to the usage model to account for self selectivity bias (McFadden et al., 1985; Mannering and Winston, 1985; Train, 1986; Hensher, *et al.*, (1992). Empirically, however, the selectivity corrections applied in utilization models to account for endogeneity bias have not had substantial effects on estimation results (Train, 1986; Hensher, 1992). More complex models, including direct joint estimation of choice and usage, are planned for future research.

2. DATA

The data are from a 1993 survey conducted using geographically stratified pure random digit dialing. The survey, covering most of urbanized California excluding San Diego County, was composed of three distinct components, as described in Brownstone, *et al.* (1994) and Golob, *et al.* (1995). An initial computer-aided telephone interview (CATI) collected information on household structure, vehicle inventory, housing characteristics, employment data, commuting for all workers and students, and information about the intended next vehicle transaction. These CATI data were then used to produce a customized mail-out questionnaire which asked detailed questions about each household member's commuting and vehicle usage. The mail-out questionnaire also contained two SP (stated preference) vehicle type choice experiments for each household, the responses to which were collected in the third part of the survey, a follow-on CATI survey.

Each of the SP experiments described three hypothetical vehicles in terms of attributes such as body type, fuel type, refueling range, purchase price, etc. These hypothetical vehicles included both alternative-fuel and gasoline vehicles. Attribute descriptions were varied according to an experimental design. For attributes related to body type and purchase price, candidate levels in the design were customized to be consistent with the types of vehicles that households indicated an interest in for their next intended vehicle purchase, so as to make the choice task more realistic and relevant. Households were asked to choose their preferred vehicle and indicate whether the chosen vehicle would replace an existing household vehicle (and if so, which one) or be added to the household fleet The questionnaire was customized so that the choices were characterized as transactions relative to the household's current vehicle holdings.

Vehicle usage SP questions followed the choice experiment. The usage questions involved asking the household to assign principal drivers to each vehicle in the new vehicle fleet (including the SP chosen vehicle), and to indicate how many miles per year the chosen vehicle was likely to be driven. The flow of the survey ensured that respondents first reported the principal drivers and usage patterns for their current vehicles before performing the SP task, allowing them to make informed judgments based on this information, as well as their own perceptions of principal driver and vehicle characteristics. Under this approach each completed survey provided both RP and SP measures of annual vehicle miles traveled that could be jointly analyzed in an appropriate model structure. The SP observations had the potential to provide data on the effect of alternative-fuel vehicle attributes on annual vehicle miles traveled.

Of the 7,387 households that completed the initial CATI survey, 66%, or 4747 households, successfully completed the mail-out portion of the survey. A comparison with Census data reveals that the sample is slightly biased toward home-owning larger households with higher incomes, and weights are being developed to balance the sample to the known population (Brownstone et al., 1994). An unweighted sample is used here.

The breakdown of the 4747 households by vehicle ownership level was: 1% zero vehicles, 34% one vehicle, 47% two vehicles, 13% three vehicles, and 5% four or more vehicles. For one-vehicle households, 75% had exactly one driver, while 25% had two or more drivers. Thus, approximately 73% of the sample households were either multi-vehicle or single-vehicle/multiple-driver, where driver allocation behavior is relevant.

The model variables are divided into three groups: (1) *behavioral* vehicle usage characteristics, capturing the ways in which households use their vehicles, (2) *physical* vehicle characteristics and (3) household structural characteristics. Vehicle usage for each household's vehicles (RP usage data) is self-reported in terms of "How many miles per year is this vehicle driven?" It would be more accurate to calculate annual usage from vehicle odometer readings one year apart, but such data are not available in a cross-sectional survey. Vehicle usage for the hypothetical future vehicles (SP usage data) was collected through a series of questions asking how many miles the vehicle chosen in a choice experiment would be driven each week, and who in the household would typically use the chosen vehicle to commute to work or school.

Because these models are being used in conjunction with a forecasting system, the household variables selected were limited to those produced by the available demographic forecasting model. The variable "Mean age of household heads," was computed as the mean of the ages of mates in spousal-like households, or the age of the single parent or person who can be identified as the major income-earner. The dummy variable "Household heads are retired" is set to one if one or both household heads are retired and neither household head is employed or temporarily unemployed; it is possible that another person, perhaps a grown child, is employed in such a household.

Separate models are developed for single-vehicle households and multi-vehicle households. The sample sizes are: 2,260 single-vehicle observations, comprised of households currently holding one vehicle (RP data) and households ending up with one vehicle after the SP choice task; 3,150 multi-vehicle observations, comprised of households currently holding two or more vehicles (RP data) and households ending up with two or more vehicles after the SP choice task. (In addition, a model of third-vehicle use was developed using a sample of 445 households with 3 or more vehicles, but this model is not reported here.) Each of the samples consists of households with known type and vintage of their current vehicle (single-vehicle sample), or no missing data on the newest two vehicles in their fleet (multi-vehicle sample). It is also required that there be no missing data on the age, sex and employment status of the principal drivers of each of these vehicles.

3. SPECIFICATION

3.1. Partition of the Variables into Endogenous and Exogenous Sets

A distinguishing feature of this research is the endogenous treatment of driver allocation behavior, as noted previously. In order to avoid omitted-variables bias, vehicle miles of travel (VMT) is specified as a function of principal driver characteristics in addition to exogenous household and vehicle type characteristics. However, principal driver characteristics are also specified as a function of the exogenous variables. This allows the principal driver characteristics, for which no exogenous forecasts are available, to be replaced by their predictors in the final forecasting equations.

There are four endogenous variables for each vehicle. These are listed in Table 1. In the multi-vehicle case, household vehicles are arranged such that the newest of the vehicles is defined as "vehicle 1," described by the first four endogenous variables and the first group of vehicle-type exogenous variables. The second-newest vehicle is defined to be "vehicle 2," and it is described by the last four endogenous variables and the last group of vehicle-type variables. If two vehicles are of the same vintage, the order of listing by the respondent is preserved.

Table 1: Endogenous Variables for Each Vehicle

Variable	Acronym
Natural log of vehicle miles traveled per year	Ln(VMT)
Age of principal driver in years	Driver Age
Gender of principal driver (0 = male; 1 = female)	Driver Gender
Employment status of principal driver (1 = working)	Driver Empl. St.

The exogenous variables in each model are divided into two blocks: physical vehicle characteristics and household characteristics. The first block, listed in Table 2, is made up of 16 physical vehicle characteristics for each vehicle.

The second block of exogenous variables is comprised of up to eleven household characteristics. This list is reproduced with associated acronyms for further reference in Table 3. The dummy variable for three or more vehicles is used only in the two-vehicle model. These variables, together with the driver characteristic variables listed in Table 1, were selected on the basis of published vehicle usage model results (Mannering, 1983; Hensher, 1985; Mannering and Winston, 1985; Hensher and Smith, 1986; Train, 1986; Golob, 1990; Hensher, at al., 1992).

Table 2: Exogenous Variables for Each Vehicle

Variable	Acronym
Vehicle Age (in years from 1993)	Vehicle Age
Vehicle class dummies (12); base = luxury class	
Mini class	Type: Mini
Subcompact car class	Type: Subcompact
Compact car class	Type: Compact
Mid-size or full-size car class	Type: Mid-size
Full-size (standard) car class	Type: Full-size
Sports car	Type: Sports Car
Compact pickup truck	Type: Small Truck
Full-size (standard) pickup truck	Type: Std. Truck
Minivan (compact van)	Type: Minivan
Full-size (standard) van	Type: Van
Compact sport utility vehicle	Type: Compact SUV.
Full-size (standard) sport utility vehicle	Type: Full-size SUV.
Operating cost per mile (in cents)	Operating Cost
Electric vehicle (dummy)	Electric Vehicle
Range between refueling in miles	Range

Table 3: Exogenous Variables - Household Characteristics

Variable	Acronym
Household membership variables	
Total number of household members > 15 years	No. 16+ Yr. Olds
Number of children in household aged 0 to 5	No. less than 5 Yrs. Old
Number of household members aged 16-20	No. 16-20 Yr. Olds
Total number of children (all ages) in household	Total children
Household is a couple (dummy)	Couple HH
Household income less than $31,000 (dummy)	Income < $31k
Household income more than $60,000 (dummy)	Income > $60k
Household head(s) are retired (dummy)	Retired HH
Mean age of household heads	Ave. Age of Heads
Total Number of workers in household	No. Heads Working
Household has three or more vehicles (dummy)	3+ Vehicle HH

3.2. The Structural Equation Model Form

The standard structural equations model (without latent variables) is given by

$$\mathbf{y} = \mathrm{B}\mathbf{y} + \Gamma\mathbf{x} + \zeta \tag{1}$$

where $\mathbf{y}$ is an mx1 column vector of endogenous variables, and $\mathbf{x}$ is an nx1 column vector of exogenous variables. The structural parameters are the elements of the matrices:

$$\underset{(m\times m)}{\mathrm{B}} = \text{matrix of causal links among the endogenous variables,}$$

and

$$\underset{(m\times n)}{\Gamma} = \text{matrix of direct causal (regression) effects from the n exogenous variables to the m endogenous variables.}$$

The . error-term parameters are the elements of the variance-covariance matrix:

$$\underset{(m\times m)}{\Psi} = E(\zeta\zeta') = \text{symmetric variance-covariance matrix of the unexplained, or unique, portions of the endogenous variables.}$$

For identification of system (1), **B** must be chosen such that **(I-B)** is non-singular, where **I** denotes the identity matrix of dimension m.

The total effects of the endogenous variables on each other are given by

$$\mathrm{H} = (\mathbf{I} - \mathrm{B})^{-1} - \mathbf{I}\,. \tag{2}$$

And the total effects of the exogenous variables on the endogenous variables in a structural equations model of this type are given by

$$\mathrm{T} = (\mathbf{I} - \mathrm{B})^{-1}\,\Gamma, \tag{3}$$

which are the parameters of the reduced-form equations.

3.3. Division of the Problem into Separate Models

Comparisons of sample sizes to the number of variables and potential number of parameters revealed that separate models could be developed for single-vehicle households and for two-vehicle households. However, the number of households with three or more vehicles was insufficient for the development of a dedicated three-vehicle model. The alternative was to expand the two-vehicle model to cover households with two *or more* vehicles, and to add a third-vehicle model for households with three or more vehicles. It might be preferable to model utilization of three vehicles simultaneously, but the expansion of the present structural equation system to 12 endogenous variables and up to 58 exogenous variables (16 for each vehicle plus 10 household variables) is infeasible with the present data. The use of a Third-Vehicle Model (not reported here), with only four endogenous and 26 exogenous variables, is a pragmatic solution to the problem.

The Two-Vehicle Model, covering the two newest vehicles in multi-vehicle households, is the most complex, and its specification is described here in detail. The Single-Vehicle Model is a simplification of the Two-Vehicle Model.

3.4. Specification of the Two-Vehicle Model

Each model specification can be subdivided into: endogenous effects given by the **B** matrix in equation system (1), exogenous effects (the Γ matrix), and error-term variance-covariances (the Ψ matrix). This specification is based on structure of the RP (revealed-preference) utilization model developed in Golob, *et al.* (1995), but the present model exhibits additional features particularly related to joint SP-RP estimation.

The postulated causal relationships among the endogenous variables are depicted in Table 4. There are two types of direct effects: within-vehicle effects and between-vehicle effects.

Table 4: Two-Vehicle Model
Postulated Direct Effects Between Endogenous Variables
(positive effects denoted by +, negative effects by -)

	Influencing Variable							
Influenced variable	Ln (VMT_1)	Driver Age_1	Driver $Gender_1$	Driver Empl. St_1	Ln (VMT_2)	Driver Age_2	Driver $Gender_2$	Driver Empl. St_2
Ln (VMT_1)		$\beta_{1,2}$ (-)	$\beta_{1,3}$ (-)	$\beta_{1,4}$ (+)				
Driver Age_1								$\beta_{2,8}$ (-)
Driver $Gender_1$		$\beta_{3,2}$ (-)					$\beta_{3,7}$ (-)	
Driver Empl. St_1		$\beta_{4,2}$ (-)	$\beta_{4,3}$ (-)					$\beta_{4,8}$ (-)
Ln (VMT_2)						$\beta_{5,6}=\beta_{1,2}$	$\beta_{5,7}=\beta_{1,3}$	$\beta_{5,8}=\beta_{1,4}$
Driver Age_2				$\beta_{6,4}=\beta_{2,8}$				
Driver $Gender_2$			$\beta_{7,3}=\beta_{3,7}$			$\beta_{7,6}=\beta_{3,2}$		
Driver Empl. St_2				$\beta_{8,4}= \beta_{4,8}$		$\beta_{8,6}=\beta_{4,2}$	$\beta_{8,7}=\beta_{4,3}$	

The *within-vehicle* effects are those in the upper left-hand (first vehicle) and lower right-hand (second vehicle) quadrants of the **B** matrix. Each of these effects is expected to be identical for the two vehicles, and equality restrictions are specified for corresponding pairs of **B**-matrix parameters. Use is postulated to be less for vehicles primarily driven by older persons ($\beta_{1,2} = \beta_{5,6}$), and women ($\beta_{1,3} = \beta_{5,7}$), and use is postulated to be greater for vehicles primarily driven by employed persons ($\beta_{1,3} = \beta_{5,8}$). Male principal drivers are more likely to be employed ($\beta_{4,3} = \beta_{8,7}$), as are younger principal drivers ($\beta_{4,2} = \beta_{8,6}$), and older drivers are expected to be male ($\beta_{3,2} = \beta_{7,6}$). An important feature of this specification is that, for each of the two household vehicles, VMT is postulated to be a function of all three of the principal driver variables. Thus, while driver allocation is endogenous, VMT is specified as a function of driver characteristics.

The postulated *cross-vehicle* effects are those in the lower left-hand and upper right-hand quadrants of the **B** matrix of Table 4. These reciprocal effects capture the relationships between the characteristics of the principal drivers of the two vehicles. We expect strong negative relationships between principal-driver genders and employment status in the between-vehicle effects, and this is operationalized by specifying equated pairs of reciprocal effects ($\beta_{3,7} = \beta_{7,3}$) and ($\beta_{4,8} = \beta_{8,4}$). In addition, we expect that, if the driver of either vehicle is employed, the driver of the other vehicle is likely to be younger than otherwise expected ($\beta_{2,8} = \beta_{6,4}$). Two-vehicle households with no employed drivers are likely to contain retired (and therefore older) members. The postulated model is parsimonious in that it has only nine free parameters in the **B** matrix, representing nine pairs of equated direct effects.

The postulated structure of the vehicle-characteristic exogenous effects is depicted in Table 5. The *vehicle-type* effects specified in the exogenous variable structure were developed by considering vehicle usage stereotypes. For example, there are typically more male principal drivers of compact and full-size pickup trucks; subcompact cars might have younger principal drivers; and minivans are likely to be driven by females. Logically, older vehicles and higher operating cost vehicles should be driven less, *ceteris paribus*.

The major restrictions applied in specifying these exogenous vehicle-type influences are that the effects be the same for the two vehicles. It is a straightforward procedure to subsequently test whether the model can be significantly improved by releasing these cross-vehicle parameter equality restrictions. It is also quite possible that the characteristics of the first vehicle can affect the VMT and principal driver characteristics of the second vehicle, and conversely. The model was initially specified by setting all such cross-vehicle effects to zero. Tests were then conducted to ascertain whether cross-vehicle effects significantly improved model fit.

Table 5: Two-Vehicle Model
Postulated Direct Effects from the Exogenous Variables

<table>
<tr><td></td><td colspan="8">Endogenous Variable</td></tr>
<tr><td>Exogenous Variable</td><td>Ln (VMT_1)</td><td>Driver Age_1</td><td>Driver $Gender_1$</td><td>Driver Empl. St_1</td><td>Ln (VMT_2)</td><td>Driver Age_2</td><td>Driver $Gender_2$</td><td>Driver Empl. St_2</td></tr>
<tr><td>Vehicle Age_1</td><td colspan="4" rowspan="5">First vehicle: effects of vehicle characteristics on VMT and driver allocation (dense sub-matrix, equated with 2nd vehicle effects)</td><td colspan="4" rowspan="5">Effects of characteristics of the first vehicle on VMT and driver allocation of the second (sparse sub-matrix, initially specified to be null)</td></tr>
<tr><td>Vehicle $Classes_1$</td></tr>
<tr><td>Operating $Cost_1$</td></tr>
<tr><td>Electric $Vehicle_1$</td></tr>
<tr><td>$Range_1$</td></tr>
<tr><td>Vehicle Age_2</td><td colspan="4" rowspan="5">Effects of characteristics of the second vehicle on VMT & driver allocation of the first vehicle (sparse sub-matrix, initially specified to be null)</td><td colspan="4" rowspan="5">Second vehicle: effects of vehicle characteristics on VMT and driver allocation (dense sub-matrix, equated with 1st vehicle effects)</td></tr>
<tr><td>Vehicle $Classes_2$</td></tr>
<tr><td>Operating $Cost_2$</td></tr>
<tr><td>Electric $Vehicle_2$</td></tr>
<tr><td>$Range_2$</td></tr>
<tr><td>No. 16-20 Yr. Olds</td><td colspan="4" rowspan="6">Effects of household characteristics on VMT and principal driver allocation of 1st vehicle (equated across vehicles)</td><td colspan="4" rowspan="6">Effects of household characteristics on VMT and principal driver allocation of 2nd vehicle (equated across vehicles)</td></tr>
<tr><td>No. of 16+ Yr. Olds</td></tr>
<tr><td>No. 1-5 Yr. Olds</td></tr>
<tr><td>Total No. of Kids</td></tr>
<tr><td>Income > $60k</td></tr>
<tr><td>Ave. Age of Heads
3+ Vehicle HH</td></tr>
</table>

Examples of direct *household effects* to be tested include: Principal drivers in households with more workers and in high-income households are more likely to be employed. Usage is higher in households with more children and in high-income households. Principal drivers are younger in households with young children. Drivers in retired households are older and are less likely to be employed (although some drivers in retired households, such as adult children living with their parents, could be employed). Finally, households with three or more vehicles have lower levels of usage on their first and second vehicles, all else held constant. The default restriction on all of these postulated household influences involves equating the corresponding effects on the two vehicles, and then testing whether the relaxation of each equality results in a significant model improvement.

The final specification step involves the error-term variance-covariance matrix Ψ. If the unique (error) component of any one of the four endogenous variables of the first vehicle is correlated with the unique component of the corresponding variable for the second vehicle, then we should find statistically significant coefficients for the Ψ matrix terms $\psi_{5,1}$, $\psi_{6,2}$, $\psi_{7,3}$, or $\psi_{8,7}$. That is, if what is not explained about a variable for one vehicle is correlated with what is not explained about the same variable for the other vehicle, these sub-diagonal parameters should be found to be significant. The freely estimated main-diagonal variances of the Ψ matrix produce R^2 values:

$$R^2 = (s_{i,i} - \psi_{i,i}) / s_{i,i} \tag{4}$$

where $s_{i,i}$ is the sample variance of endogenous variable i.

4. ESTIMATION METHOD

Structural equations systems of this type can be generally estimated using methods of moments (also known as variance analysis methods). These methods proceed by defining the sample variance-covariance matrix of the combined set of endogenous and exogenous variables, partitioned with the endogenous variables first:

$$\mathbf{S} = \begin{bmatrix} S_{yy} & S_{yx} \\ S'_{yx} & S_{xx} \end{bmatrix}, \tag{5}$$

where $\mathbf{S}_{yy}$ denotes the variance-covariance matrix of the endogenous variables, $\mathbf{S}_{yx}$ denotes the covariance matrix between the endogenous and exogenous variables, and $\mathbf{S}_{xx}$ denotes the variance-covariance matrix of the exogenous variables. In the Two-Vehicle Model, there are 8 endogenous variables and 38 exogenous variables, so $\mathbf{S}$ is a (46 by 46) symmetric matrix. It can be easily shown using matrix algebra that the corresponding variance-covariance matrix replicated by model system (1), denoted by

$$\Sigma = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma'_{yx} & \Sigma_{xx} \end{bmatrix}, \tag{6}$$

is:

$$\Sigma_{yy} = (I-B)^{-1}(\Gamma S_{xx}\Gamma' + \Psi)\left((I-B)^{-1}\right)', \tag{7}$$

$$\Sigma_{yx} = (I-B)^{-1}\Gamma S_{xx}, \tag{8}$$

and $\Sigma_{xx} = S_{xx}$ is taken as given. The structural equation system is estimated using the variance-analysis normal-theory maximum likelihood method (Bollen, 1989). The fitting function for structural equations maximum likelihood (ML) estimation is

$$F_{ML} = \mathrm{Log}|\Sigma(\theta)| - \mathrm{Log}|S| + \mathrm{tr}[S\Sigma(\theta)] - (p+q) \ , \tag{9}$$

where $\Sigma(\theta)$ represents Σ (equations 6-8) implied by the vector of model parameters, θ. This fitting function is (-2/n) times the log of the likelihood function that **S** is observed if $\Sigma(\theta)$ is the true multivariate normal variance-covariance matrix. Minimization of $\mathbf{F}_{ML}$ is equivalent to maximization of the likelihood function. Under the assumptions of multivariate normality and the model being correctly specified, $n\mathbf{F}_{ML}$ is chi-square distributed, providing a test of model rejection and criteria for testing hierarchical models. Function (9) is minimized in the LISREL8 program using a modified Fletcher-Powell algorithm (Jöreskog and Sörbom, 1993a).

Note that if we fit the unrestricted reduced form by regressing each endogenous variable on all of the exogenous variables, then $\Sigma(\theta)$ would exactly equal **S** and $\mathbf{F}_{ML}$ would be zero. Alternatively, we would get exactly the same result if we only imposed enough restrictions on the underlying parameters, θ, to just identify the system. All of our models impose more restrictions than are necessary to identify the system. It turns out that $n\mathbf{F}_{ML}$ is simply the likelihood ratio test statistic for the null hypothesis that these over-identifying parameter restrictions are consistent with our observed data.

Because four of the eight endogenous variables are dichotomous, the structural error terms corresponding to these dichotomous variables will have unequal variances. Although the coefficient estimates will still be consistent, the estimates of parameter standard errors and the overall model chi-square goodness-of-fit will be inconsistent (Bentler and Bonett, 1980). Consistent estimates can be generated using the asymptotically distribution-free weighted least squares method (Browne, 1982, 1984), but this requires a much larger sample size. (The rule-of-thumb is that the sample size must be at least three times greater than the number of free entries in the asymptotic variance-covariance matrix of the correlation matrix, i.e., the fourth order moments; with 36 variables, this requires approximately 3,250 observations.) However, Amemiya (1981) shows that for moderate sample sizes, the weighted least squares method may produce worse results than the unweighted maximum likelihood estimates used here. In any case, the coefficient estimates are still consistent. and they have been shown to be fairly robust (Boomsma, 1983).

5. RESULTS: TWO-VEHICLE MODEL

5.1. Model Fit and Structure

The Two-Vehicle Model fits well according to the standard goodness-of-fit criteria. The likelihood ratio test statistic associated with the null hypothesis that the estimated model is consistent with the observed sample variance-covariance matrix is 210.5 with 237 degrees of freedom, corresponding to a probability value of 0.892. Thus, the model *cannot* be rejected at the p = . 10 level. The estimated R^2 value for VMT of the first (newest) vehicle is 0.115, and that of the second (oldest) vehicle is 0.131. As expected, significant positive error-term covariances were found between the VMT's of the two vehicles (*t-statistic* = 11.3), between principal driver ages (*t-statistic* = 8.4), and between principal driver genders (*t-statistic* = 15.2).

The estimated direct effects between endogenous variables are listed with their *t*-statistics in Table 6. This endogenous variable structure model is basically in accordance with the hypotheses depicted in Table 4. All six of the within-vehicle endogenous-variable effects postulated for each vehicle were found to be statistically significant and five of the six effects are equal across the two vehicles. The three postulated cross-vehicle effects were also found to be significant and symmetric.

Table 6: Two-Vehicle Model
Estimated Direct Effects Between Endogenous Variables (*t*-statistics in parentheses)
(Coefficients that are restricted to be equal for the two vehicles are shown in **bold**)

	Influencing Variable							
Influenced variable	Ln (VMT_1)	Driver Age_1	Driver $Gender_1$	Driver Empl. St_1	Ln (VMT_2)	Driver Age_2	Driver $Gender_2$	Driver Empl. St_2
Ln (VMT_1)		**-0.0043** **(-4.05)**	**-0.131** **(-6.52)**	**0.179** **(4.54)**			0.0797 (2.70)	
Driver Age_1								**-2.81** **(-5.85)**
Driver $Gender_1$		-0.0051 (-7.78)					**-0.693** **(-21.6)**	
Driver Empl. St_1		**-0.0065** **(-10.3)**	**-0.103** **(-11.3)**					**-0.140** **(-15.9)**
Ln (VMT_2)		-0.0028 (-2.12)				**-0.0043** **(-4.05)**	**-0.131** **(-6.52)**	**0.179** **(4.54)**
Driver Age_2				**-2.81** **(-5.85)**				
Driver $Gender_2$			**-0.693** **(-21.6)**	0.506 ((3.72)		-0.0036 (-6.46)		
Driver Empl. St_2				**-0.140** **(-15.9)**		**-0.0065** **(-10.3)**	**-0.103** **(-11.3)**	

Three additional cross-vehicle effects were found to be necessary for good model fit. These are identified by the highlighted cells in Table 6: (1) If the driver of the first vehicle is older, use of the second vehicle is less than otherwise expected (effect $\beta_{5,2}$), an effect which further links the influence of driver-age on VMT for the two vehicles. (2) If the driver of the second vehicle is female, use of the first vehicle is greater than expected (effect $\beta_{5,2}$). And (3) if the driver of the first-vehicle is employed, the driver of the second vehicle is more likely to be female (effect $\beta_{7,4}$). The last two effects are consistent with the travel and activity patterns in many households in which there are working male heads and non-working female heads who bear the primary child care and home management responsibilities (Robinson, 1977; Pas, 1984; Townsend, 1987; Golob and McNally, 1995).

5.2. Total Effects

The total effects of the endogenous variables on the two vehicle usage variables are listed in Table 7. For simplicity, only the total effects on the two VMT variables, elements $\eta_{1,j}$ and $\eta_{5,j}$ (j = 1 to 8) of matrix H defined in equation system (2), are shown. Results show that *driver age* has a significant effect on vehicle usage that is uniform for the two vehicles; if *either* driver is younger, *both* the first and second vehicles are likely to be used more. In contrast, the *gender* and *employment status* effects are consistent and reciprocal across the two vehicles. If the principal driver of either vehicle is female, that vehicle is driven less and the other vehicle is driven more, and if either driver is employed, that vehicle is driven more, and the other vehicle is driven less, *ceteris paribus*. These reciprocal pairs of effects are generally strongest for the driver's own vehicle.

Table 7: Two-Vehicle Model
Total Effects of the Other Endogenous Variables on the Two Usage Variables

	Influenced Variable			
Endogenous Variable	Ln (VMT_1)		Ln (VMT_2)	
	Total effect	*t*-statistic	Total effect	*t*-statistic
Driver Age_1	-0.00358	-3.36	-0.00363	-2.79
Driver $Gender_1$	-0.40013	-7.13	0.19975	5.33
Driver Empl. St_1	0.20385	5.03	-0.02565	-3.10
Driver Age_2	-0.00116	-4.72	-0.00362	-2.91
Driver $Gender_2$	0.35896	5.60	-0.28460	-6.24
Driver Empl. St_2	-0.02098	-3.26	0.15278	3.84

The total effects of the exogenous variables on the usage endogenous variables are listed in Table 8. These are the coefficients of the reduced-form equations for two of the eight endogenous variables, which are given by matrix equation (2).

Table 8: Two-Vehicle Model
Total Effects of the Exogenous Variables on the Two Vehicle Usage Variables

Exogenous Variable	Total Effects			
	Ln (VMT_1)		Ln (VMT_2)	
	Total effect	*t*-statistic	Total effect	*t*-statistic
Vehicle Age_1	-0.01301	-2.98	-0.00095	-2.48
$Type_1$: Mini car	-0.23091	-5.33	-0.00043	-0.10
$Type_1$: Subcompact	0.01675	3.33	0.01983	3.49
$Type_1$: Compact car	0.08289	2.41	0.00710	2.39
$Type_1$: Mid-size car	-0.01416	-2.69	-0.06500	-1.57
$Type_1$: Full-size car	-0.08733	-1.57	-0.00872	-2.32
$Type_1$: Sports car	0.03025	3.97	0.01494	2.44
$Type_1$: Small truck	0.07210	8.35	-0.03599	-6.52
$Type_1$: Std. Truck	0.08037	8.37	-0.04012	-6.53
$Type_1$: Minivan	0.12686	2.44	0.01668	4.12
$Type_1$: Std. Van	0.02095	2.05	-0.01046	-2.01
$Type_1$: Small SUV.	0.23267	4.33	-0.02145	-3.25
$Type_1$: Std. SUV.	0.07242	6.74	-0.02177	-2.80
Operating $Cost_1$	-0.00057	-1.36	-0.00058	-1.31
Electric $Vehicle_1$	-0.25025	-2.51	0.09420	1.25
$Range_1$	0.00153	3.30	0	- -
Vehicle Age_2	0.00443	1.14	-0.03372	-9.34
$Type_2$: Mini car	0.00323	3.19	-0.16784	-3.29
$Type_2$: Subcompact	-0.00290	-0.59	0.02421	3.99
$Type_2$: Compact car	0.00140	2.26	0.00436	1.93
$Type_2$: Mid-size car	0.01319	2.46	-0.01784	-3.21
$Type_2$: Full-size car	0	- -	0	- -
$Type_2$: Sports car	-0.09826	-2.12	0.02713	3.63
$Type_2$: Small truck	-0.06246	-6.43	0.05819	6.99
$Type_2$: Std. Truck	-0.08374	-6.86	0.05863	6.70
$Type_2$: Minivan	0	- -	0	- -
$Type_2$: Std. Van	-0.01879	-2.01	0.01490	2.02
$Type_2$: Small SUV.	-0.03829	-4.68	0.12405	2.17
$Type_2$: Std. SUV.	-0.05326	-5.30	0.05463	5.84
Operating $Cost_2$	0.00427	0.626	-0.00860	-1.19
Electric $Vehicle_2$	0	- -	-0.22579	-1.20
$Range_2$	-0.00096	-1.53	0.00072	0.81
No. of 16-20 Yr. Olds	0.00956	4.80	0.00822	1.55
No. less than 5 Yrs. Old	-0.00917	-3.41	0.04667	1.95
Total no. of Kids	0.03060	3.30	0.03662	3.88
Income>$60k	0.11339	4.13	0.08506	3.10
Retired HH	-0.05129	-4.84	-0.04452	-4.19
Ave. Age of Heads	-0.00350	-4.24	-0.00545	-5.88
No. Heads Working	0.11234	5.51	0.10618	5.20
3+ Vehicle HH	0	- -	-0.04580	-1.59

The total effects of *vehicle age* on VMT are strongest for the second vehicle, but the effects are consistent for both vehicles: The older a vehicle is, the less it is used, *ceteris paribus*. Also, the older the first vehicle is, the less the other vehicle is used as well. The forecasting implication of this is reduced usage of the household fleet over time if no vehicle transactions occur. If household structure, income and employment do not change, the reduction in the fleet VMT will be further accentuated through the negative total effect on usage of driver age. This implies that households wishing to accommodate new travel demand are more likely to replace a vehicle with a newer one; while households with declining travel demand are more likely to hold on to their existing vehicles.

The total effects of *operating cost* are imprecisely estimated, but the signs of the within-vehicle effects are as expected. Also, a higher operating cost for the second vehicle implies a shift of usage from the second vehicle to the first vehicle, but the coefficients in the reduced-form equations have relatively high standard errors.

The availability of the SP usage data yielded information about the effect of a limited range vehicle on annual VMT that would not otherwise be available from the RP responses alone. The effects of the *electric vehicle* (EV) dummy variable on VMT are potentially important for pollution and energy policies. If either of the first two vehicles in multi-vehicle households is a future EV, the model results imply that the EV will be driven less, *ceteris paribus*. Moreover, if the EV is the newest (first) vehicle in the household, the second vehicle will be driven more than otherwise expected. Thus, this model captures a shift in usage from EV's to conventional fuel vehicles, somewhat mitigating the emissions gains of the electricity versus conventional fuels. The magnitude of this cross-vehicle substitution effect can be assessed by using this utilization model for forecasting in combination with demographic, vehicle transaction, and vehicle type choice models (Brownstone, *at al.*, 1994).

The *range* variable also captures a reduced VMT effect for all limited-range vehicles (potentially including dedicated compressed natural gas vehicles in addition to EV's).

For limited-range second vehicles, there is also a shift in usage from the second vehicle to the first vehicle.

The *number of household members between 16 and 20 years old* has a positive influence on VMT of both the first and second vehicle. However, the *number of drivers* in the household has negative effects on VMT of both vehicles, possibly indicating a shift of usage toward third and fourth vehicles in the household. The *number of children 1 to 5 years old* positively influences VMT mostly of the second vehicle, while the *total number of children* positively influences VMT of <u>both</u> the first and second vehicles. The *income* effect has the expected sign, but, as in the case of *average age of the heads*, the effects are imprecisely estimated. Finally, as expected, the presence of *three or more household vehicles* reduces VMT of both the first and second vehicles.

5.3. Scenarios of Changes in VMT Implied by the Total Effects

The endogenous variables are expressed in terms of the natural logarithms of VMT, so the natural exponent of each reduced-form equation coefficient represents a multiplicative factor applied to the endogenous VMT variable in question. That is, $\exp(\delta_j \tau_{1,j})$ and $\exp(\delta_j \tau_{5,j})$ express multipliers of VMT for vehicles 1 and 2, respectively, where the T matrix of total exogenous effects is defined in equation system (3) and δ_j is the level of change in the jth exogenous variable. Some selected VMT multiplier effects are listed in Table 9 and 10.

Each scenario listed in Tables 9 and 10 assumes that all factors not defined in the scenario remain constant. In the case of vehicle replacements, this includes the vehicle type class and operating cost. However, to provide realism, when vehicles are assumed to be replaced with identical vehicles with different ranges or fuels, it is assumed that the replacement vehicle is one year newer.

Of all of the model predictions computed in Table 9, the most substantial effects are those attributable to *vehicle range* and the *electric vehicle (EV) designator.* In the case of the first (newer) vehicle, a reduction in range of 150 miles reduces VMT by a factor of 0.81, but there is no effect on VMT of the second vehicle. In the case of the second vehicle, a similar reduction in range of 150 miles reduces VMT by a factor of only 0.93, but first-vehicle VMT is predicted to increase by a factor of 1.15. The weaker second-vehicle range effect is partially due to an offsetting stronger second-vehicle age effect. Combining reduced range with the EV effect, the model predicts that if the first vehicle is an EV with 100 miles range, VMT will reduce by a factor of 0.58, and second-vehicle VMT will increase by a factor of 1.10. If the second vehicle is an EV with 100 miles range, VMT on this vehicle will reduce by a factor of 0.70, but there will be more of a shift to usage of the first vehicle, with first-vehicle VMT increasing by a factor of 1.24.

Table 9: Two-Vehicle Model
Exponentiated Total Effects on VMT of Selected Changes in Vehicle Characteristics

Exogenous change	Multiplier Effect on Veh. 1 VMT	Multiplier Effect on Veh. 2 VMT
Vehicle Age_1 (in years)		
vehicle ages one year	0.98	0.99
replace with same type veh. 1 year newer	1.01	1.00
replace with same type veh. 5 years newer	1.07	1.01
Electric $Vehicle_1$ (EV_1)		
$Range_1$ (in miles)		
replace 300 mi. veh. with 200 mi. non-EV, 1 yr. newer	0.87	1.00
replace 300 mi. veh. with 150 mi. non-EV, 1 yr. newer	0.81	1.00
replace 300 mi. veh. with 100 mi. EV, 1 yr. newer	0.58	1.10
replace 300 mi. veh. with 75 mi. EV. 1 yr. newer	0.56	1.10
Vehicle Age_2 (in years)		
vehicle ages one year	1.00	0.97
replace with same type veh. 1 year newer	0.99	1.03
replace with same type veh. 5 years newer	0.98	1.18
Electric $Vehicle_2$ (EV_2)		
$Range_2$ (in miles)		
replace 300 mi. veh. with 200 mi. non-EV, 1 yr. newer	1.10	0.96
replace 300 mi. veh. with 150 mi. non-EV, 1 yr. newer	1.15	0.93
replace 300 mi. veh. with 100 mi. EV, 1 yr. newer	1.21	0.72
replace 300 mi. veh. with 75 mi. EV, 1 yr. newer	1.24	0.70

Given these results, we regard the SP data as providing useful improvements to the quality of our VMT forecasts for future alternative-fuel vehicles, and especially for electric vehicles. However, there were also some possible limitations inherent in the data from the SP experiment, which was mainly focused on the issue of vehicle choice. Respondents were apparently able to reflect the general effect of limited range electric vehicles on usage patterns through both the allocation of the vehicle and some adjustments to VMT. More subtle effects on utilization due to other attributes such as limited fuel availability (e.g., away from home recharging for electric vehicles, or smaller numbers of stations for natural gas vehicles) or differences in fuel operating costs may not have been as easily captured using this experimental format.. This could have resulted in an over-estimation of range and EV effects, and an underestimation of the effects of, e.g., improved operating costs. In fact, the coefficient on operating cost in our current model is rather modest, implying that the range and EV scenario results would not be substantially changed by imposing accompanying realistic changes in operating costs. Invalidating this result would require additional research.

In contrast to the range effects, the *vehicle aging* effects are weaker for the first (newer) vehicle than for the second (older) vehicle. If the newest vehicle in the household is replaced with a vehicle that is identical in type, operating cost, range, and fuel, but is five years newer, the model predicts that VMT for that vehicle will increase by approximately seven percent, with very little effect on VMT of the second vehicle. However, if the second vehicle is replaced with a vehicle that is identical in type, operating cost, range, and fuel, but is five years newer, the model predicts that VMT for that vehicle will increase by approximately eighteen percent, and VMT of the first vehicle will slightly decrease (by about two percent).

predicted changes in VMT associated with the scenarios related to household characteristics are listed in Table 10. The influences related to the *number of children in the household* are smaller in magnitude than expected, but the usage behavior appears consistent with conventional notions of first- and second-vehicles. For example, a new child places more pressure on the use of the second vehicle, the one that is less likely to be used for commuting.

In contrast, the predicted effects of *income* and *the number of household heads working* are relatively strong, especially in combination. The joint impact of an additional worker and a higher household income is a predicted increase of twenty-five percent in usage for the first household vehicle, approximately half of which is attributable to an income effect; usage of the second vehicle increases by a slightly lower twenty-one percent. If one working head retires *and* income drops below the high-income cut-off, the model predicts that VMT of the first and second vehicles will be reduced by the factors 0.85 and 0.88, respectively. If both household heads quit working due to retirement, the predicted change in VMT is only ten percent for the first vehicle, providing that household income remains above (or below) the high-income cut-off. Finally, the presence of a third household vehicle has a modest influence on VMT of the second vehicle.

Table 10: Two-Vehicle Model
Exponentiated Total Effects on VMT of Selected Changes in Household Characteristics

Exogenous change	Multiplier Effect on Veh. 1 VMT	Multiplier Effect on Veh. 2 VMT
Number of 16-20 Yr. Olds		
child passes 16th birthday, no other changes	1.01	1.02
Children		
birth of child, no other changes	1.02	1.09
19 yr. old child moves out of home, no other changes	0.96	0.96
Income and Employment Status		
income rises above $60k, no other changes	1.12	1.09
+1 head working *and* income rises above $60k	1.25	1.21
1 head working: retires, income drops below $60k	0.85	0.88
2 heads retire at same time, income stays above $60k	0.90	0.92
Ownership of a Third Vehicle		
household adds third vehicle	1.00	0.96
household disposes of third vehicle	1.00	1.05

6. RESULTS: THE SINGLE-VEHICLE MODEL

6.1. Model Fit and Final Structure

The structure of the Single-Vehicle Model is also basically in accordance with the structural hypotheses. This Model fits extremely well according to all goodness-of-fit criteria, the chi-square statistic being 41.82 with 49 degrees of freedom, corresponding to a probability value of 0.757. The model *cannot* be rejected at the p = .10 level. The estimated R^2 value for VMT is 0.173. No significant error-term covariances were found between any pairs of the four endogenous variables.

The endogenous variable structure determined to be optimal in the Single-Vehicle Model is similar to the within-vehicle structure found for the Two-Vehicle Model (the structure depicted in the upper-left-hand and lower-right-hand quadrants of the **B** matrix shown in Table 6). The only difference was that an additional direct effect was found between principal driver gender and age: female principal drivers of a vehicle in a single-vehicle household are younger than otherwise expected, *ceteris paribus*.

6.2. Total Effects

The total effects of the endogenous principal-driver variables on VMT for the Single-Vehicle Model are listed in Table 11. As in the multi-vehicle case, VMT is higher for younger, male, employed drivers, but the gender and employment status effects are relatively weaker for single-vehicle households.

Table 11: Single-Vehicle Model
Total Effects of the Other Endogenous Variables on Vehicle Usage

Endogenous Variable	Total Effect on Ln(VMT)	
	Total effect	*t*-statistic
Driver Age	-0.00396	-1.00
Driver Gender	-0.08037	-2.30
Driver Empl. Status	0.11671	2.46

Finally, the total exogenous effects on VMT for the Single-Vehicle Model are listed in Table 12. Once again, these effects are similar to those found for multi-vehicle households, with some exceptions. Usage patterns are consistent for eight types of vehicles, but sportscars, minivans, standard sport utility vehicles, and full-size cars exhibit different usage patterns in single-vehicle, versus multi-vehicle, households

Table 12: Single-Vehicle Model
Total Effects of the Exogenous Variables on Vehicle Usage

Exogenous Variable	Total effect	*t*-statistic
Vehicle Age	-0.01574	-4.14
Type: Mini car	-0.27808	-5.97
Type: Subcompact	0.09798	1.82
Type: Compact car	0.12140	2.26
Type: Mid-size car	-0.00259	-0.91
Type: Full-size car	0.00639	1.38
Type: Sports car	-0.00706	-1.46
Type: Small truck	0.26612	2.64
Type: Std. Truck	0.52883	3.26
Type: Minivan	0.45711	3.86
Type: Std. Van	0.34705	1.64
Type: Small SUV.	0.31306	2.98
Type: Std. SUV.	0.00000	0.00
Operating Cost	-0.01223	-1.31
Electric Vehicle	-0.15136	-1.58
Range	0.00138	3.62
No. 16-20 Yr. Olds	0.04246	2.58
No. 16+ Yr. Olds	0.03455	0.93
No. less than 5 Yrs. Old	0.12448	2.31
Total no. of Kids	-0.11225	-3.78
Income < $31k	-0.19112	-5.19
Income > $60k	0.10970	1.90
Couple HH	0.00833	1.69
Retired HH	-0.02178	-1.67
Ave. Age of Heads	-0.01071	-8.46
No. Heads Working	0.05588	2.21

Regarding *alternative-fuel vehicles*, the negative EV effect and the positive effect of range on VMT are consistent between single-vehicle and multi-vehicle households.

7. A FORECASTING METHOD THAT PRESERVES HETEROGENEITY

This model is being applied in a dynamic microsimulation forecasting system (Brownstone, *et al.*, 1994), where a sociodemographic transition model and vehicle transactions models are being used to forecast changes in households' sociodemographic structure and composition of the vehicle fleets. The usage model is then exercised to forecast VMT for both the before- and after-situations for the household. The calculated change in forecasts is then applied as a percentage change to the actual base level of usage for the household in the before-situation.

Even if the dynamic sociodemographic model predicts no change in household characteristics (household composition, employment status, or income), and the vehicle transactions model predicts no vehicle transactions for the household for the period in question, the present usage model will in general predict changes in VMT. This will be due to aging of the household heads, aging of the vehicles, and possible changes in the age categories of household members, particularly children.

The most effective application of the usage models in a micro-simulation forecasting system uses a "pivot" approach (described below), rather than the traditional approach of using the expected value from a linear model. The pivot approach preserves heterogeneity across households. Heterogeneity due to spatial and lifestyle factors is to be expected; some households drive more miles per year than the model would predict while others drive fewer miles per year than the model would predict.

By using the residual difference between observed and predicted VMT for each household vehicle in the model estimation data set, we can develop household/vehicle-specific multipliers that can be used during forecasting. Such multipliers take the form:

$$\delta_i = VMT_{i,0}^{Obs} / VMT_{i,0}^{Pred}. \tag{10}$$

where i denotes the ith vehicle, and 0 denotes the "base year" of the forecast, which corresponds to the original estimation sample. A new predicted VMT is computed for each forecasting period, and then "pivoted" by using the multiplier. These multipliers capture effects due to heterogeneity that might be missing from the model, and preserves them in the forecasts. One difficulty with this approach is that vehicle transactions will occur during the course of a forecast. Appropriate rules have been developed for re-assigning multipliers after a vehicle transaction, and these are being used in our microsimulation forecasting system.

8. CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

The structural elegance of the models and their statistical fit to the sample data provide support to our modeling approach. Moreover, the correspondence between pure RP results (Golob, *et al.*, 1995) and the present SP-RP results is encouraging. We are also encouraged by the advantages associated with a jointly estimated RP-SP model that simultaneously captures the endogenous effects of vehicle re-allocation along with perceived changes in utilization associated with electric vehicle characteristics. These effects are not available from RP data alone. The approach automatically produces estimates that are consistently scaled, and yields reduced form equations that are convenient for forecasting utilization of alternative fuel vehicles.*?*

However, the SP questions in the 1993 household survey from which these data were extracted are primarily focused on the issue of vehicle choice, and are potentially limited in capturing the full range of effects on usage attributable to fuel availability, peak and off-peak recharge costs for EV's, cargo capacity, performance, and other vehicle and fuel-system characteristics that might distinguish future vehicles. These issues are being pursued through a second household survey, conducted in 1994 that contained a different vehicle usage SP protocol. When the 1994 data are available, the robustness of the present model results can be assessed, and hopefully the model can be extended.

Potential selectivity bias can be accounted for in this usage model by linking the model to a discrete type-choice model (e.g., Ren, et al., 1995), and adding into the structural equation system a correction term variable involving a transformation of the household's predicted type vehicle choice probabilities (McFadden et al., 1985; Mannering and Winston, 1985; Train, 1986; Hensher, et al., 1992). It is doubtful that such a correction term would have a pronounced effect on the results.

The known biases in the normal-theory maximum likelihood estimation method applied to dichotomous endogenous variables are concentrated on coefficient standard errors and overall goodness-of-fit criteria. The fit of the model is not in question, and hypothesis testing is subordinate to forecasting capability in this research. However, it would be possible to use unbiased generally weighted least squares estimation (Browne, 1982, 1984), as implemented in LISREL8 with PRE-LIS2 (Jöreskog and Sörbom, 1993b), with a significantly increased sample size.

9. ACKNOWLEDGMENTS

This research is part of a project to develop a model system to forecast demand for clean-fuel vehicles in California, sponsored by Southern California Edison Company and the California Energy Commission. The authors wish to thank their colleagues, who have made the research reported here possible: Jane Torous, Weiping Ren and Seyoung Kim of the University of California, Irvine; Ryuichi Kitamura of the University of California, Davis and Kyoto University; and Mark Bradley of the Bradley Research and Consulting Group. Ernest Morales, Richard Rice, and Debbie Brodt of Southern California Edison Co., and Gary Occhiuzzo and Mike Jaske of the California Energy Commission have also worked diligently on behalf of the project. The authors are solely responsible for the views expressed here and for any remaining errors.

REFERENCES

Amemiya, T. (1981). Qualitative response models, a survey. *Journal of Economic Literature*, 19: 481-536.

Bentler, P.M. and Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88: 558-606.

Berkovec, J. (1985). Forecasting automobile demand using disaggregate choice models. *Transportation Research,* 19B: 315-329.

Bollen, K.A. (1989). *Structural Equations with Latent Variables.* Wiley, New York.

Boomsma, A. (1983). *On the Robustness of LISREL (Maximum Likelihood Estimation) against Small Sample Sizes and Nonnormality.* Amsterdam: Sociometric Research Foundation.

Browne, M.W. (1982). Covariance structures. In D.M. Hawkins, ed., *Topics in Multivariate Analysis:* 72-141. Cambridge: Cambridge University Press.

Browne, M.W. (1984). Asymptotic distribution free methods in analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37: 62-83.

Brownstone, D., D.S. Bunch and T.F. Golob (1994). A demand forecasting system for clean-fuel vehicles. Presented at the OECD International Conference "Towards Clean Transport: Fuel Efficient and Clean Motor Vehicles," Mexico City, March 28-30.

Brownstone, D., D.S. Bunch, T.F. Golob and W. Ren (1995). A personal vehicle transactions choice model for use in forecasting demand for future alternative-fuel vehicles. *Research in Transportation Economics*, in press.

Bunch, D.S., M. Bradley, T.F. Golob, R. Kitamura and G.P. Occhiuzzo (1993). Demand for clean-fuel personal vehicles in California: A discrete-choice stated preference study. *Transportation Research,* 27A: 237-253.

Bunch, D.S., D. Brownstone and T.F. Golob (1995). A dynamic forecasting system for vehicle markets with clean-fuel vehicles. To be presented at Seventh World Conference on Transport Research, Sydney, July 16-21.

Golob, T.F. (1990). The dynamics of travel time expenditures and car ownership decisions. *Transportation Research*, 24A: 443-463.

Golob, T.F., S. Kim and W. Ren (1995). A structural model of vehicle use in multi-vehicle households. Presented at Annual Meeting of Transportation Research Board, Washington, DC, January 23-28.

Golob, T.F., R. Kitamura, M. Bradley and D.S. Bunch (1993). Predicting the market penetration of electric and clean-fuel vehicles. *The Science of the Total Environment,* 134: 371-381.

Golob, T.F. and M.G. McNally (1995). A Model of household interactions in activity participation and the derived demand for travel. Presented at the *Conference on Activity based Approaches: Activity Scheduling and the Analysis of Activity Patterns,* May 25-28, 1995, Eindhoven University of Technology, the Netherlands.

Hensher, D.A. (1985). An econometric model of vehicle use in the household sector. *Transportation Research*, 19B: 303-314.

Hensher, D.A. and V. Le Plastrier (1985). Towards a dynamic discrete-choice model of household automobile fleet size and composition. *Transportation Research,* 19B: 481-495.

Hensher, D.A. and Manefield, T. (1982). A structured-logit model of automobile acquisition and type choice. *Proceedings of the 7th Australian Transport Research Forum, Hobart, Tasmania:* 625-654.

Hensher, D.A. and N.C. Smith (1986). A structural model of the use of automobiles by households. *Transport Reviews,* 6: 87-111.

Hensher, D.A., N.C. Smith, F.W. Milthorpe and P.O. Bernard (1992). *Dimensions of Automobile Use: A Longitudinal Study of Automobile Ownership and Use.* Amsterdam: North Holland.

Hocherman, I., J. Prashker, and M. Ben-Akiva (1983). Estimation and use of dynamic transaction models of automobile ownership, *Transportation Research Record No. 944:* 134-141.

Jöreskog, K.G. and D. Sörbom (1993a). *LISREL 8 User's Reference Guide.* Scientific Software, Chicago.

Jöreskog, K.G. and D. Sörbom (1993b). *PRELIS 2 User's Reference Guide.* Scientific Software, Chicago.

Kazimi, C. (1994). Hazard models of changing household demographics. Working Paper UCI-ITS-WP-94-10, Institute of Transportation Studies, University of California, Irvine, December.

Kazimi, C. and D. Brownstone (1995). Competing risk hazard models for demographic transitions. To be presented at the Econometric Society World Congress, Tokyo, Aug. 27-29.

Lave, C.A. and K. Train (1979). A disaggregate model of auto-type choice. *Transportation Research,* 13A: 1-9.

Mannering, F. (1983). An econometric analysis of vehicle use in multi-vehicle households. *Transportation Research,* 17A: 183-189.

Mannering, F. and C. Winston (1985). Dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics,* 16: 215-236.

McFadden, D. L., C. Winston and A. Borsch-Supan (1985). Joint estimation of freight transportation decisions under non-random sampling. In A.F. Daugherty, ed. *Analytical Studies in Transport Economics:* 137-160. Cambridge University Press, Cambridge.

McCarthy, P.S. and R. Tay (1989). Consumer valuation of new car attributes: an econometric analysis of the demand for domestic and Japanese and Western European imports. *Transportation Research,* 23A: 367-375.

Pas, E.I. (1984). The effects of selected sociodemographic characteristics on daily travel-activity behavior. *Environment and Planning A*, 16: 571-581.

Robinson, J.P. (1977). *How Americans use time: A Social-Psychological Analysis of Everyday Behavior*, Praeger, New York.

Townsend, T.A. (1987). The Effects of Household Characteristics on the Multi-day Time Allocations and Travel - Activity Patterns of Households and Their Members. Unpublished Ph.D. Dissertation, Northwestern University, Evanston, IL.

Train, K. (1986). *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand.* Cambridge: MIT Press.

van Wissen, L.J. and T.F. Golob (1992). A dynamic model of car fuel type choice and mobility. *Transportation Research*, 26B: 77-96.