

1. Report No. FHWA/TX-00/0-1838-5	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle VMT MIX MODELING FOR MOBILE SOURCE EMISSIONS FORECASTING: FORMULATION AND EMPIRICAL APPLICATION		5. Report Date May 2000	
		6. Performing Organization Code	
7. Author(s) Chandra R. Bhat and Harikesh S. Nair		8. Performing Organization Report No. 0-1838-5	
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 3208 Red River, Suite 200 Austin, TX 78705-2650		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. 0-1838	
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Transfer Section/Construction Division P.O. Box 5080 Austin, TX 78763-5080		13. Type of Report and Period Covered Research Report (9/99-5/00)	
		14. Sponsoring Agency Code	
15. Supplementary Notes Project conducted in cooperation with the Federal Highway Administration and the Texas Department of Transportation.			
16. Abstract The purpose of the current report is to propose and implement a methodology for obtaining improved link-specific vehicle miles of travel (VMT) mix values compared to those obtained from existent methods. Specifically, the research is developing a fractional split model that predicts the VMT mix on links as a function of the functional roadway classification of the link, the physical attributes of the link, the operating conditions on the link, and the attributes of the traffic analysis zone in which the link lies.			
17. Key Words Nonattainment areas, ozone, transportation control measures, emissions modeling, travel demand modeling		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.	
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 40	22. Price

**VMT MIX MODELING FOR MOBILE SOURCE EMISSIONS FORECASTING:
FORMULATION AND EMPIRICAL APPLICATION**

by

Chandra R. Bhat and Harikesh S. Nair

Research Report Number 0-1838-5

Research Project 0-1838

Transportation Control Measure Effectiveness in Ozone Nonattainment Areas

Conducted for the

TEXAS DEPARTMENT OF TRANSPORTATION

in cooperation with the

U.S. DEPARTMENT OF TRANSPORTATION

FEDERAL HIGHWAY ADMINISTRATION

by the

CENTER FOR TRANSPORTATION RESEARCH

Bureau of Engineering Research

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2000

DISCLAIMERS

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration or the Texas Department of Transportation. This report does not constitute a standard, specification, or regulation.

There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine, manufacture, design or composition of matter, or any new and useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES

Principal Investigator

Dr. Chandra R. Bhat

ACKNOWLEDGMENTS

The authors acknowledge the support of TxDOT Project Director William Knowles of TxDOT Austin District Construction's Division, George Reeves of TxDOT Dallas District, Wayne Young of TNRCC TxDOT Environmental Affairs Division and Carol Nixon of TxDOT Houston District. The authors also acknowledge Arnold Breedan of TxDOT and Ken Cervenka, Ken Kirkpatrick, Gustavo Baez, Mahmoud Ahmadi and Chris Klaus of NCTCOG for providing documents and data.

Research performed in cooperation with the Texas Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.

TABLE OF CONTENTS

1. BACKGROUND AND SIGNIFICANCE OF WORK.....	1
2. STATE OF THE ART/PRACTICE IN VMT MIX DETERMINATION	3
3. FRACTIONAL SPLIT MODEL STRUCTURE	5
3.1. General Background.....	5
3.2. Quasi-Likelihood Estimation	7
4. DATA PREPARATION	10
4.1. Data Sources.....	10
4.2. Data Assembly	12
5. EMPIRICAL ANALYSIS	13
5.1. Sample Description	13
5.2. Fractional Split Model Results	15
6. MODEL APPLICATION	19
7. MODEL EVALUATION.....	22
7.1. VMT Mix Performance Evaluation.....	23
7.2. Emissions Performance Evaluation.....	25
8. CONCLUSIONS.....	27
ACKNOWLEDGMENTS.....	29
REFERENCES.....	31

LIST OF TABLES

Table 1: Fractional Split of Vehicle Types	13
Table 2: VMT Fractional Split Model Estimation Results	16
Table 3: TxDOT Vehicle Count Vehicle Type to MOBILE Vehicle Type Conversion Factors.....	21
Table 4: VMT Mix Performance Evaluation	24
Table 5: Emissions Performance Evaluation	26

1. BACKGROUND AND SIGNIFICANCE OF WORK

The integration of transportation planning and air quality planning is important for mobile-source emissions estimation. The Environmental Protection Agency (EPA) requires the use of MOBILE5 for such emissions estimation for all areas except California, which uses the EMFAC7F model.

The emissions factor models (MOBILE5 and EMFAC7F) require several traffic-related inputs, including travel speeds, vehicle miles of travel (VMT), on-road operating conditions (operating mode of vehicles, environmental conditions, and existence of inspection/maintenance programs), vehicle age distribution by vehicle class in the area of analysis, and vehicle mileage accumulation rates by vehicle class. These inputs are used to calculate emissions factors (in grams per mile of vehicle travel for each pollutant) for eight different vehicle classes. The vehicle class-specific emissions factors are then applied to the VMT accumulated by each of the eight vehicle classes and are finally aggregated across all vehicle classes to obtain the total vehicular emissions.

The level of detail at which the emissions analysis is conducted varies substantially among metropolitan regions. But the EPA requires that metropolitan planning areas rated as serious or higher in nonattainment designation for ozone and carbon monoxide estimate their mobile-source emissions using network-based transportation models. The planning organizations in these areas generally conduct their emissions analysis at an individual link level. This level involves the estimation of volumes and speeds on each network link in the metropolitan area from travel demand models, followed by the computation of link-specific emissions based on a) link VMT, b) vehicle speed on the link, c) the vehicle class-specific

emissions factors, and d) VMT mix fractions in the eight vehicle classes. Of these, the link VMT and link speeds are obtained directly from the network-based travel demand models. The vehicle class-specific emissions factors are obtained from the emissions-factor models based on the various inputs listed previously. The VMT fraction by vehicle class (referred to as VMT mix in the MOBILE5 model) is a supplementary traffic-related parameter that is to be provided by the analyst.

The emissions factors for each of the three pollutants, carbon monoxide (CO), volatile organic compounds (VOC), and oxides of nitrogen (NO_x), vary widely among the different vehicle classes. Consequently, the emissions analysis is very sensitive to VMT mix. For example, at high temperatures, a 2.8% change in the heavy-duty gas vehicle (HDGV) mix causes about a 10% change in the CO emissions rate, and a 4.8% change in the HDGV mix leads to about a 10% shift in the VOC emissions rate (see Chatterjee *et al.* [1], page 45). It is, therefore, important to provide accurate VMT mix values at the individual link level (see *NCHRP Research Results Digest* [2], which identifies improvement in VMT mix modeling as an area of research that would be particularly beneficial for emissions modeling). The purpose of the current report is to propose and implement a methodology for obtaining improved link-specific VMT mix values compared to those obtained from existent methods. Specifically, the research is developing a fractional split model that predicts the VMT mix on links as a function of the functional roadway classification of the link, the physical attributes of the link, the operating conditions on the link, and the attributes of the traffic analysis zone in which the link lies.

The rest of this paper is structured as follows. The next section reviews the state of the art/practice in VMT mix modeling. Section 3 presents the structure of, and the estimation

technique for, the fractional split model used in the current report. Section 4 provides an overview of the data sources used in the study and discusses issues related to data preparation. Section 5 presents empirical results. Section 6 focuses on model application. Section 7 discusses model evaluation issues. The final section concludes the paper.

2. STATE OF THE ART/PRACTICE IN VMT MIX DETERMINATION

The emission-factor models require the total VMT accumulated on a road to be by split across eight vehicle classes. The vehicle classes are based on the size and weight of vehicles as well as on the type of fuel used. The eight vehicle classes are light-duty gasoline vehicle (LDGV), light-duty gasoline truck, type 1 (LDGT1), light-duty gasoline truck, type 2 (LDGV2), heavy-duty gasoline vehicle (HDGV), light-duty diesel vehicle (LDDV), light-duty diesel truck (LDDT), heavy-duty diesel vehicle (HDDV), and motorcycle (MC).

The current practice in many metropolitan areas is to accept the aggregate default VMT mix computed by MOBILE5 and to apply this mix to all network links. The default VMT mix is based on national data reflecting the proportion of travel by each vehicle type in urban areas.

Another approach adopted by some metropolitan agencies is using 24-hour local vehicle-classification counts (rather than MOBILE5 default values) to determine VMT mix. The agencies then apply various factors to convert the vehicle types in traffic counts to the eight MOBILE5 vehicle classes. EPA recommends that local agencies adopt this approach because the MOBILE5 default values may not be reflective of the local traffic vehicle mix. In this local vehicle count-based approach, the VMT mix is typically stratified by the functional classification of roadways to accommodate variations across roadway classes. However,

since most counts are conducted only on higher roadway classes (such as interstates and major arterials), there is inadequate information for agencies to comprehensively capture variations in VMT mix by roadway class. Values of VMT mix obtained for the higher roadway classes are applied (sometimes after *ad hoc* adjustments) to the lower roadway classes (such as minor arterials, collectors, and local roads).

A problem with the state of the art/practice discussed above for VMT mix determination is that it applies aggregate-level values across links in the road network in a region. In an analysis of VMT mix from 477 different count sites in the U.S., Chatterjee *et al.* [1] found substantial variation in VMT mix across the sites, emphasizing the need for local determination of VMT mix values (rather than using MOBILE default values). The same study also indicates substantial variation in VMT mix even after controlling for roadway class at any given site, underscoring the need to consider explanatory factors other than roadway class in local VMT mix analyses.

The discussion above motivates the research documented in this paper. Specifically, we formulate and estimate a fractional split model that determines the VMT mix ratio as a function of several informative variables, which include the physical attributes of links (such as number of lanes and whether the link is a divided road or not), the operating characteristics of links (such as link speed), aggregate area-type characterizations of the traffic survey zone in which the link lies (such as urban, suburban, and rural), and the land-use attributes of the zone (such as retail acreage and manufacturing/warehouse acreage in the zone). Such a model will facilitate accurate VMT mix computations at a fine level of geographic resolution.

3. FRACTIONAL SPLIT MODEL STRUCTURE

3.1. General Background

Fractional response-dependent variables arise naturally in many transportation analyses and other analytical contexts. Examples of such variables include the proportion of freight tonnage for a commodity group moving from an origin to each of several destinations, the proportion of intercity trips made by each of several travel modes, time spent by an individual in one of several activity types (such as shopping, social-recreational, or personal business), and (as in the present analysis) the fraction of VMT accrued by each vehicle class. A characteristic of all these analyses is that the variable of interest is in the form of fractions. The sum of the fractions across all categories of the variable is equal to one, and each fraction is bounded between zero and one. In addition, one or more of the fractions may take the boundary values of zero or one. In the following discussion, we present the fractional split model structure in the context of VMT mix analysis.

Mathematically, let y_{qi} be the fraction of VMT accrued by vehicle type i ($i=1,2,\dots,I$) on link q . Let this fraction be a function of a vector x_q of relevant explanatory variables associated with attributes of the link and the traffic analysis zone in which the link lies. A common approach to analyzing fractional dependent variables is to model the log-odds ratio as a linear function (for example, see Bhat and Misra [3]):

$$E\left(\log \frac{y_{qi}}{y_{q1}}\right) = \beta_i' x_q, i \neq 1, \quad (1)$$

where β_i is a parameter vector to be estimated for each i (except for a base category which needs to be normalized to zero for identification; in the above equation, the first category is arbitrarily assigned to be the base category; *i.e.*, $\beta_i = \mathbf{0}$, where $\mathbf{0}$ is the null vector of the appropriate size). y_{qt} is the VMT fraction accrued by the first vehicle type. If some parameters in the β_i vectors are equal across categories, such restrictions can be imposed by jointly estimating all β_i vectors after appropriate data structuring (see Bhat and Misra, [3]).

The specification in Equation 1 is attractive since the transformed dependent variable in the regression is unbounded and can take values anywhere on the real line as y_{qt} varies between 0 and 1. Thus, a linear regression is appropriate. However, as pointed out by Papke and Wooldridge [4], the specification has at least two major problems. First, the dependent variable is undefined when the fraction of VMT in a vehicle class is 0/1. If the numbers of vehicle class observation combinations for which the boundary conditions prevail is small, arbitrary small adjustments may be made prior to the computation of the log-odds ratio without significantly affecting the estimated parameters. However, if there are several vehicle class-observation combinations for which the boundary conditions prevail, the adjustments can have a substantial impact on estimation. In our analysis, the fractional VMT for some vehicle types (such as buses and trucks) is 0 for a large percentage of observations (*i.e.*, links) for which vehicle classification counts are available.¹

¹If the dependent variable represents proportions from a fixed number of groups with known group sizes, suitable adjustments have been proposed in the econometric literature (see Maddala [5]; page 30). However, the corresponding Berkson's minimum chi-squared estimation method is not applicable when the fractions arise naturally in analysis settings (such as the current VMT mix setting) rather than as a result of the discrete grouping of disaggregate observations.

A second problem with the specification in Equation 1 is that, even if the econometric specification in Equation 1 is appropriate and well defined, one cannot obtain $E(y_{qi} | x_q)$ (which is of primary interest for VMT fraction forecasting) without making additional assumptions about the distribution of the residuals, $u_{qi} = \log[y_{qi} / y_{q1}] - \beta_i' x_q$ ($i = 2, 3, \dots, I$). If a distribution is assumed or estimated, then $E(y_{qi} | x_q)$ may be computed by first obtaining the conditional expected value (on residuals) for each fraction and then unconditioning out the residuals by integrating over the distribution of the assumed or estimated distribution for the residuals (see Bhat [6] for an application of this method). However, this approach is either nonrobust (if an incorrect parametric distribution is assumed) or cumbersome (if a nonparametric distribution for the residuals is estimated). Also, the integration in this method will involve as many dimensions as there are vehicle types, and this can become tedious.

3.2 Quasi-Likelihood Estimation

The model we propose here for VMT mix modeling is a polychotomous extension of the binary fractional split model proposed by Papke and Wooldridge [4]. The approach does not need any *ad hoc* adjustment for boundary values of the dependent variable fractions, and it directly specifies a model for $E(y_{qi} | x_q)$. At the same time, the approach is easy to implement and is robust, since we make no assumptions about the distribution of y_{qI} conditional on x_q . The focus is on consistent estimation of the parameters appearing in the conditional mean specification $E(y_{qi} | x_q)$ and on consistent, asymptotically robust estimation of the standard errors of the conditional mean parameters.

Consider the following econometric specification:

$$E(y_{qi} | x_q) = G_i(\beta, x_q), \quad 0 < G_i(\cdot) < 1, \quad \sum_{j=1}^I G_j(\cdot) = 1, \quad \text{where } \beta = (\beta_2', \beta_3', \dots, \beta_I)'. \quad (2)$$

$G_i(\cdot)$ ($i=1,2,\dots,I$) in the above equation is a predetermined function, and the properties specified for it above ensure that the predicted fractional VMT in each vehicle class for any link will lie in the interval (0,1) and will sum to 1 across vehicle classes. The econometric model in Equation 2 is well defined even if y_{qi} takes on the value of 0 or 1 with positive probability. The reader will note that the specification above does not make any assumption about the true underlying conditional distribution of y_{qi} given x_q . This distribution is considered unknown and can have any underlying structure.

The β parameter vector in the conditional mean model of Equation 2 is estimated by maximizing a likelihood function associated with a family of probability distributions that does not necessarily contain the true unknown distribution. The label “quasi-likelihood estimation” is used for such estimations (see *Gourieroux et al.* [7]). Specifically, we use the multinomial log-likelihood function in the quasi-estimation:

$$\mathcal{L}_q(\beta) = \sum_{i=1}^I y_{qi} \log G_i(\beta, x_q). \quad (3)$$

The multinomial quasi-likelihood estimator used above belongs more generally to the linear exponential family (LEF). *Gourieroux et al.* [7] prove the strong consistency and asymptotic normality of the parameter estimator of the conditional mean (*i.e.*, the elements of the β vector) obtained by maximizing $\sum_q \mathcal{L}_q(\beta)$, as long as (and if and only if) $\mathcal{L}_q(\beta)$ belongs to the LEF (see also *Wooldridge* [8]). This is a very strong result, since it is based only on the correct specification of the conditional mean function of Equation 2. The result holds

irrespective of the true distribution of y_{qi} conditional on x_q . Of course, if we are able to correctly specify this true distribution, we can maximize the true likelihood function in order to obtain an estimator more efficient than the quasi-likelihood estimator used here. However, the disadvantage of this alternative approach (compared to the quasi-approach) is that the resulting “true-likelihood” estimator is inconsistent under an incorrect assumption for the true distribution.

Within the family of LEF-based quasi-likelihood estimators, asymptotic efficiency can be achieved if the functional form of the true conditional second-order moment (*i.e.*, variance) of y_{qi} given x_q is known. This case is unlikely. We prefer to base our inference only on the conditional mean specification of Equation 2 and to propose consistent and asymptotically robust inferences for the conditional mean parameter vector β . As indicated by Papke and Wooldridge [4], this can be achieved by computing the asymptotic variance-covariance matrix of β as $H^{-1}\Delta H^{-1}$, where H is the Hessian and Δ is the cross-product matrix of the gradients (H and Δ are evaluated at the estimated parameter values).

A final model structure issue concerns the specification of the functional form for G_i in the conditional mean specification of Equation 2. We use a multinomial logit functional form for G_i since this structure is easy to program and implement. In this structure, we write:

$$G_i(\beta, x_q) = \frac{e^{\beta_i x_q}}{\sum_{j=1}^I e^{\beta_j x_q}}, \text{ where } (\beta_2', \beta_3', \dots, \beta_T)'. \quad (4)$$

4. DATA PREPARATION

4.1. Data Sources

Several data sources are used in the current analysis. These sources include the following: a) vehicle classification counts conducted in the Dallas-Fort Worth area by the Texas Department of Transportation (TxDOT) Regional Planning Organization (RPO) and TxDOT's Transportation Planning and Programming Division; b) 1996 GIS-based road network file for the Dallas-Fort Worth area; c) zonal level land-use characteristics file of the Dallas-Fort Worth area; and d) 1996 GIS-based Dallas-Fort Worth zonal coverage file. The latter three data files were obtained from the North Central Texas Council of Governments (NCTCOG). Each of the four data sources is discussed briefly in the following paragraphs.

The TxDOT vehicle classification counts used in the analysis were conducted at several fixed stations in the Dallas-Fort Worth area during the periods from 1977 to 1987 and 1983 to 1993. The counts covered all the functional roadway classes and a mixture of land uses, with intent to obtain a sample that is representative of the VMT mix in the region. The counts were recorded using a manual count board from 6 a.m. to 10 p.m. on weekdays in the same month every year. The 16-hour counts were expanded to a 24-hour period to allow researchers to obtain the 24-hour vehicle classification counts. These 24-hour counts form the basis for computing the VMT mix on links. The counts distinguish among the following vehicle types: automobiles, pickups and vans (PUV), sports utility vehicles (SUV), combination trucks (two axles, three axles, four axles, and six axles), buses (two axles and three axles) and motorcycles (including all two-wheelers). The counts separate trucks by the number of axles, but we combined them for the current analysis because of the frequent occurrence of zero-counts in several axle categories. We also combined two-axle and three-

axle buses into a single bus category for the same reason. The fraction of counts of each vehicle type represents the VMT mix at the individual link level and is the dependent variable in the current analysis.

The 1996 GIS road network file includes information on the characteristics of each link in the Dallas-Fort Worth metropolitan planning area. The metropolitan area (MA) includes about a 4,980 sq mi area with more than 45,000 unique roadway links to represent the roadway network. The MA covers the existing urbanized area and the contiguous area expected to be urbanized by the year 2020. It includes all of Collin, Dallas, Denton, Rockwall, and Tarrant Counties and portions of Ellis, Johnson, Kaufman, and Parker Counties. The link attributes available in the network file include length of the link, traffic direction, functional classification, number of lanes, free speed, capacity, and whether the link is divided.

The zonal level land-use characteristics file of the Dallas-Fort Worth area contains land-use data at the level of the traffic survey zone used by NCTCOG for its travel demand modeling purposes. There are about 6,000 traffic survey zones in the Dallas-Fort Worth MA. The land-use data for each zone include total land area and acreage in several individual land-use purposes (such as in manufacturing and warehousing; in retail, hotel, and motel; in institutional buildings like churches, government, museums, schools, and hospitals; and in airport runways and terminals).

The 1996 GIS-based Dallas-Fort Worth zonal coverage file provides the spatial configuration of the traffic survey zones in the Dallas-Fort Worth MA.

4.2. Data Assembly

The objective of the data-assembly steps was to append the appropriate link and zonal characteristics to each link observation. To accomplish this, we first spatially overlaid the 1996 GIS road network file and the 1996 GIS zonal coverage file. Next, each link at which vehicle counts were recorded was manually queried in the network database using the name of the street and the names of the cross streets at the end nodes. Once the link at which a count was made was spatially located in the GIS road network coverage, its identifying number in the network file was extracted. Also, the traffic analysis zone that spatially contains the link of interest was identified from the GIS zonal coverage. Using these link and zonal identifier fields, the relevant link and land-use characteristics were mapped to each vehicle count observation.

The raw TxDOT vehicle classification counts included 370 observations of link vehicle counts, of which only 244 observations could be geo-coded in the manner discussed above. These 244 link-count observations constituted the final sample for analysis. The vehicle type distribution in this final sample was almost the same as the vehicle type distribution in the raw data.²

²The reader will note that even if the aggregate sample VMT mix does not reflect the actual aggregate VMT mix in a region, the estimated model parameters will still be consistent except for the category-specific constants. This is because of the multinomial logit structure of Equation 4. Consistent values of the category-specific constants can be obtained in a straightforward fashion if the aggregate vehicle type distribution in the local region is known (see Ben-Akiva and Lerman [9], page 237).

5. EMPIRICAL ANALYSIS

5.1. Sample Description

The descriptive statistics of the dependent variable (*i.e.*, the fractional splits among the six vehicle types across observations) in the sample are provided in Table 1. As expected, on average, the automobile fractional split is highest, followed by the fraction of PUVs. The average ratio of SUVs is between 3 and 5 percent. However, at an individual link level, the SUV percentage is as high as 48.7 and the truck percentage is as high as 26.3. The fraction of buses and motorcycles in the vehicle stream is relatively low.

Table 1: Fractional Split of Vehicle Types

Vehicle Type	Mean Value	Std. Dev.	Minimum	Maximum
Autos	0.653	0.088	0.389	0.875
Pickups/Vans (PUV)	0.262	0.062	0.098	0.416
Sports Utility Vehs. (SUV)	0.035	0.034	0.000	0.487
Trucks	0.043	0.045	0.000	0.263
Buses	0.002	0.008	0.000	0.118
Motorcycles and Two-Wheelers (MC)	0.005	0.003	0.000	0.023

The percentage of observations for which the fractional mix of trucks, buses, and motorcycles is at or very close to the boundary value of 0 is rather high. In particular, the truck fraction is less than 0.01 for 33 percent of observations, the bus fraction is less than 0.01 for 99 percent of observations, and the motorcycle fraction is less than 0.01 for 95 percent of observations. Thus, using the log-odds analysis method (Equation 1) would be inappropriate in the current modeling context. The specification in Equation 2, which can accommodate boundary values of the dependent variable, is the appropriate approach.

Five sets of independent variables were included in the model to predict the VMT mix on links. These are a) link functional classification, b) link physical attributes, c) link free speed variables, d) degree of urbanization of the zone in which link lies, and e) zonal land-use characteristics. A number of variables within each of the five variable classes were considered in the model specification. The final set of variables and their method of inclusion in the VMT mix model was determined based on, first, a systematic process of eliminating variables found to be statistically insignificant in previous specifications, and, second, considerations of parsimony in representation. In the description below, we briefly highlight some of the characteristics of the variables in each of the five sets of variables that were retained in the final model specification.

The link functional classification identifies each link with one of five roadway classes: freeways, major arterials, minor arterials, collectors, and local or residential roads. Because the number of observations on local or residential roads was very small (only 4 out of the sample size of 244), we combined the collector and local/residential road classes into a single “collector/local” category. The sample split among the four resulting roadway classes is as follows: freeways (41.8 percent), major arterials (26.2 percent), minor arterials (13.5 percent), and collector/local links (18.5 percent).

Two link physical attributes turned out to be important determinants of the link VMT mix; these were the number of lanes and whether the link was divided. Of the links in the sample, a majority (56.6 percent) has two lanes; 10.2 percent of links have one lane, 24.2 percent have three lanes, and 9 percent have four lanes. A substantial percentage (82.4) of links are divided roads.

The link free speed varies between 9 mph and 68 mph, with a mean value of 45 mph. A direct specification with free speed as the independent variable was inferior to the specification that categorized links into one of four free speed groups: low speed (less than or equal to 30 mph), low-to-medium speed (31 to 40 mph), medium speed (41 to 55 mph), and high speed (greater than 55 mph).

The degree of urbanization of the zone in which the link lies is determined by classifying the zone as a central business district (CBD), an urban residential area, or a suburban or rural area (because the differentiation between suburban and rural areas did not affect the VMT mix, these two categories were combined). In the sample, about 5 percent of links are in a CBD area, about 40 percent in an urban area, and the remainder in a suburban or rural area.

The zonal land use variables include a) an airport presence variable indicating the presence or absence of airport runway or terminal facilities in the zone in which the link lies; b) an institution presence variable indicating the presence or absence of institutions such as churches, schools, and hospitals; c) zone acreage in retail and office space; and d) acreage in manufacturing and warehousing. In the sample, about 5 percent of links lie in a zone with airport-related infrastructure and about 49 percent of links are in zones with some land use for institutional facilities. The average zone acreage in retail and office space in the sample is 18.43 acres and the average zone acreage in manufacturing and warehousing is 31 acres.

5.2 Fractional Split Model Results

The final model specification results are provided in Table 2. The table provides estimates of the β parameter vector in Equation 4.

The link functional classification variables are introduced with the freeway class acting as the base roadway category. The results indicate an increase in the PUV and motorcycle fractions on major and minor arterials relative to freeways and to other vehicle types. The fraction of these two vehicle types, however, is highest on collector/local streets. The bus fraction is lower on minor arterials (compared to freeways and major arterials) and even lower on collector or residential streets.

Table 2: VMT Fractional Split Model Estimation Results

Variable	Parameter Estimate	t-Statistic
Link functional classification		
<u>Major arterials</u>		
PUV	0.0934	2.090
MC	0.3595	4.130
<u>Minor arterials</u>		
PUV	0.1076	1.177
Bus	-1.0570	-2.193
MC	0.3138	2.506
<u>Collector/local streets</u>		
PUV	0.2416	3.427
Bus	-1.7264	-3.140
MC	0.6679	4.460
Link physical attributes		
<u>Divided road</u>		
Truck	1.1389	3.149
Bus	-0.6862	-2.328
MC	0.3427	2.625
<u>Number of lanes</u>		
Truck	-0.1738	-2.202
Bus	-0.5230	-2.045
Link free speed variables		
<u>Low speed</u>		
PUV	-0.2903	-3.838
SUV	-0.7688	-6.499
Truck	-1.7293	-7.013
Bus	1.0436	2.610
<u>Low-to-medium speed</u>		
PUV	-0.1469	-2.297
SUV	-0.3377	-3.259
Truck	-1.8454	-11.480
Bus	0.5063	1.744

Table 2: VMT Fractional Split Model Estimation Results (continued)

Variable	Parameter Estimate	t-Statistic
Link free speed variables (continued)		
<u>Medium speed</u>		
Truck	-0.4125	-3.847
MC	0.1481	1.829
Degree of urbanization		
<u>Central Business District</u>		
PUV	-0.2919	-3.205
Truck	-1.0350	-3.409
Bus	1.7342	4.106
<u>Urban residential</u>		
PUV	-0.0918	-1.932
SUV	-0.2322	-2.395
Truck	-0.5645	-4.209
Zonal land-use variables		
<u>Airport presence</u>		
PUV	0.1823	2.211
<u>Institution presence</u>		
Auto	0.1207	2.904
<u>Acreage in office/retail space</u>		
PUV	-0.0019	-2.399
SUV	-0.0038	-1.796
Truck	-0.0165	-4.282
Bus	-0.0146	-1.937
MC	-0.0026	-2.169
<u>Acreage in manufacturing/warehousing</u>		
PUV	0.0009	3.427
SUV	0.0021	5.494
Truck	0.0067	9.632
Bus	0.0031	2.457

- Notes: 1) PUV - Pickups and vans, SUV - Sports utility vehicle, MC - Motorcycles/two-wheelers
2) The estimated constant values for each vehicle type are as follows (the auto vehicle type is the base): -0.8147 (for PUV), -2.1601 (for SUV), -2.4148 (for trucks), -4.2927 (for buses), and -5.3752 (for MC).

The results of the effect of link physical attributes indicate an increase in truck fraction and a decrease in bus fraction on divided roads. Motorcycles are also more prevalent on divided highways than are other nontruck vehicle types. The impact of the number of lanes on vehicle mix suggests a decrease in the truck and bus fractions in the vehicle fleet on links with several lanes.

The link free speed variable coefficients show fewer PUVs and SUVs as a fraction of total vehicles on low-speed links relative to buses and passenger cars and relative to medium and high-speed links. The same, though more tempered, negative trend exists for PUVs and SUVs on low to medium-speed links. Generally speaking, PUVs and SUVs are more prevalent on higher-speed links than on lower-speed links. The same is true for the truck mix in the vehicle fleet, except that this effect is much stronger for trucks than for SUVs and PUVs. The results also indicate that the bus fraction is highest on low-speed facilities and is higher on low to medium-speed facilities than on high free-speed links. Finally, the fraction of motorcycles and other two-wheelers is higher on medium-speed links than on other links.

The coefficients on the variables characterizing the degree of urbanization show a lower fraction of trucks on links in CBD and urban residential zones relative to links in suburban/rural zones. Among the nontruck vehicle types, the PUV fraction is likely to be smaller than that of the other vehicle types on CBD links, and the SUV fraction is likely to be smaller than that of other vehicle types on urban links. In addition, the bus fraction is highest on CBD links compared to other link types.

The final set of variables is the set of land-use variables. The results reveal that the proportion of PUVs is high on links in zones with airport facilities. This finding is quite reasonable because PUVs are more convenient for transporting baggage and passengers to airports. The auto proportion is high on links in zones where institutions such as churches, schools, and hospitals are present. Similarly, the auto proportion is high on links in zones with large areas allocated to retail and office space. Finally, vehicle types other than automobiles and motorcycles are likely to capture a high proportion of the VMT mix in zones with large areas invested in manufacturing plants and warehouses.

6. MODEL APPLICATION

The model results in Table 2 can be applied in forecasting mode to determine the VMT mix in the six vehicle types: autos, PUVs, SUVs, trucks, buses, and motorcycles and two-wheelers. The implementation is particularly straightforward when one uses a GIS platform. This is the method the research team is using to determine the VMT mix on each link in the Dallas-Fort Worth MA as part of an ongoing air quality-related project funded by TxDOT.

The model-predicted VMT mix in the six vehicle types has to be converted into the eight-class EPA vehicle classification for input into the MOBILE emissions-factor model. We propose an approach that combines local vehicle registration data from the Dallas-Fort Worth area with information provided in the Transportation Energy Data Handbook (Davis, 10) for this conversion, as discussed below.

The *Transportation Energy Data Handbook* estimates that 98.8 percent of passenger cars are gasoline-driven and 1.2% are diesel-driven. These splits are used to allocate the “autos” VMT mix between the LDGV (light-duty gasoline-powered vehicles) and LDDV (light-duty diesel-powered vehicles) EPA categories. Pickups and vans (PUVs, including minivans and passenger vans) and sports utility vehicles (SUVs) fall under the classification of light-duty trucks and are to be assigned among the LDGT1 (light-duty gasoline-powered trucks of gross vehicle weight less than 6,000 lb), LDGT2 (light-duty gasoline-powered trucks of gross vehicle weight between 6,000 and 8,500 lb), and LDDT (light-duty diesel-powered trucks of gross vehicle weight less than 8,500 lb) EPA vehicle types. From the *Transportation Energy Data Handbook*, we computed the gasoline-to-diesel split for light-duty trucks as 97.88 percent to 2.12 percent based on truck sale information up to 1995. This

information is used to allocate 2.12 percent of the PUV VMT mix and the SUV VMT mix to the LDDT category. The remaining 97.88 percent of gasoline-powered PUVs and SUVs are allocated between the LDGT1 and LDGT2 categories based on 1996 local vehicle registration data obtained from TxDOT for the Dallas-Fort Worth region. Since the local vehicle registration data are differentiated by county, and the split of the LDGT1 and LDGT2 categories in the registration data are quite different across counties, we developed county-specific estimates of PUV and SUV splits in the LDGT1 and LDGT2 categories (our method, of course, assumes that the split of traffic in the LDGT1 and LDGT2 categories in each county is the same as the vehicle registration split in these categories in the county; the method does not consider intercounty travel, which may lead to a differential LDGT1/LDGT2 traffic split *vis-à-vis* the registration data split in each county) .

The 1996 local vehicle registration data for the Dallas-Fort Worth area provide the gasoline-diesel splits of combination trucks (vehicles with gross weight of over 8,500 lb) by county. This information is used directly to apportion the “combination trucks” VMT mix into the HDGV (heavy-duty gasoline vehicles) and HDDV (heavy-duty diesel vehicles) EPA categories.

A “bus” vehicle type classification is not included in the 1996 local vehicle registration data. Hence, we estimated the split of buses into the gasoline-powered and diesel-powered vehicles from the *Transportation Energy Data Handbook* (20.09 percent gasoline-powered and 79.91 percent diesel-powered) and used it to allocate the “bus” VMT mix between the HDGV and HDDV EPA vehicle categories.

Finally, the model-predicted “motorcycle” VMT mix is assigned completely to the MC EPA vehicle category.

Table 3 provides the final county-specific conversion factors between the six-vehicle type classification typology used in the VMT mix modeling of the current report and the eight-vehicle type EPA classification typology.

Table 3: TxDOT Vehicle Count Vehicle Type to MOBILE Vehicle Type Conversion Factors

Dallas County

TxDOT classification	EPA MOBILE vehicle type classification							
	LDGV	LDDV	LDGT1	LDGT2	LDDT	HDGV	HDDV	MC
Autos	98.8%	1.2%	-	-	-	-	-	-
PUV	-	-	95.16	2.72	2.12%	-	-	-
SUV	-	-	95.16	2.72	2.12%	-	-	-
Trucks	-	-	-	-	-	35.43%	64.57%	-
Buses	-	-	-	-	-	20.09%	79.91%	-
Motorcycles	-	-	-	-	-	-	-	100%

Tarrant County

TxDOT classification	EPA MOBILE vehicle type classification							
	LDGV	LDDV	LDGT1	LDGT2	LDDT	HDGV	HDDV	MC
Autos	98.8%	1.2%	-	-	-	-	-	-
PUV	-	-	96.07%	1.81%	2.12%	-	-	-
SUV	-	-	96.07%	1.81%	2.12%	-	-	-
Trucks	-	-	-	-	-	39.31%	60.69%	-
Buses	-	-	-	-	-	20.09%	79.91%	-
Motorcycles	-	-	-	-	-	-	-	100%

Collin County

TxDOT classification	EPA MOBILE vehicle type classification							
	LDGV	LDDV	LDGT1	LDGT2	LDDT	HDGV	HDDV	MC
Autos	98.8%	1.2%	-	-	-	-	-	-
PUV	-	-	96.15%	1.73%	2.12%	-	-	-
SUV	-	-	96.15%	1.73%	2.12%	-	-	-
Trucks	-	-	-	-	-	44.24%	55.76%	-
Buses	-	-	-	-	-	20.09%	79.91%	-
Motorcycles	-	-	-	-	-	-	-	100%

Denton County

TxDOT classification	EPA MOBILE vehicle type classification							
	LDGV	LDDV	LDGT1	LDGT2	LDDT	HDGV	HDDV	MC
Autos	98.8%	1.2%	-	-	-	-	-	-
PUV	-	-	96.36%	1.52%	2.12%	-	-	-
SUV	-	-	96.36%	1.52%	2.12%	-	-	-
Trucks	-	-	-	-	-	43.30%	56.70%	-
Buses	-	-	-	-	-	20.09%	79.91%	-
Motorcycles	-	-	-	-	-	-	-	100%

**Table 3: TxDOT Vehicle Count Vehicle Type to MOBILE Vehicle Type Conversion Factors
(continued)**

Rockwell County

TxDOT classification	EPA MOBILE vehicle type classification							
	LDGV	LDDV	LDGT1	LDGT2	LDDT	HDGV	HDDV	MC
Autos	98.8%	1.2%	-	-	-	-	-	-
PUV	-	-	95.96%	1.92%	2.12%	-	-	-
SUV	-	-	95.96%	1.92%	2.12%	-	-	-
Trucks	-	-	-	-	-	34.24%	65.76%	-
Buses	-	-	-	-	-	20.09%	79.91%	-
Motorcycles	-	-	-	-	-	-	-	100%

There are a few things to keep in mind about the conversion factors. First, these conversion factors can be updated continually as more local information becomes available. Second, the current EPA MOBILE emissions-factor model does not distinguish between PUVs and SUVs; both of these are classified as light-duty trucks. Thus, the distinction between PUVs and SUVs in our VMT mix model is rather academic at this point. However, new versions of the MOBILE models that distinguish between emissions of PUVs and SUVs are planned. The VMT mix model proposed here can be used to provide the disaggregate input needed by these forthcoming MOBILE models.

7. MODEL EVALUATION

In this section, we evaluate the ability of the proposed model to replicate the actual VMT mix on links in the sample. We also compare the predicted emissions on each link (based on the proposed VMT mix model) with the actual emissions on that link (based on the observed link-specific VMT mix values). In addition, we compare the performance of the proposed model with that of a “default” model that uses only roadway functional classification as the controlling variable for VMT mix analysis (this is the state of the

practice in the Dallas-Fort Worth area and in other metropolitan areas that use local VMT mix values). It will be noted that the “default” model is better than using the MOBILE5 default values because it is based on local conditions. But it is restrictive compared to the proposed model in this paper because it ignores factors other than roadway functional classification that may affect VMT fractions.

7.1. VMT Mix Performance Evaluation

For each sample link observation, we have data on the actual observed VMT mix in the six vehicle types: autos, PUVs, SUVs, trucks, buses, and motorcycles. We also have available for analysis the corresponding model-predicted VMT mix in these six vehicle types. The evaluation of the proximity of estimated and actual VMT mixes on links is based on three criteria: the mean absolute error (MAE), the mean percentage absolute error (MPAE), and a pseudo- R^2 value. The MAE is computed for each vehicle type as the average of the absolute difference between the model-predicted and actual VMT fractions for that vehicle type across link observations.

The MPAE is computed as the link-averaged absolute difference between the model-predicted and actual VMT fractions as a percentage of the actual VMT fraction. The pseudo- R^2 measure is an overall model-fit measure computed as shown below:

$$R^2 = \frac{\sum_q \sum_i (\hat{y}_{qi} - \bar{y}_i)^2}{\sum_q \sum_i (y_{qi} - \bar{y}_i)^2}, \quad (5)$$

In Equation 5 above, y_{qi} is the actual fraction of VMT accrued by vehicle type i on link q , \hat{y}_{qi} is the model-predicted fraction, and \bar{y}_i is the area-wide average VMT (from Table 1) for

vehicle type i . The denominator in Equation 5 is the variation in the actual link VMT mix values around the mean areawide VMT mix value, summed across all vehicle types and links. The numerator represents the variation explained by the model. Thus, the pseudo- R^2 measure may be viewed as the fraction of total variation in VMT mix explained by the model. The measure varies between 0 and 1.

Table 4 provides the measures of fit (MAE, MPAE, and pseudo- R^2) for the proposed model and a “default” model that uses only roadway functional classification as the controlling variable. We do not compute an MPAE for buses and motorcycles in Table 4 because the actual VMT fractions for these vehicles are extremely small, leading to substantially high MPAE values by construction.

Table 4: VMT Mix Performance Evaluation

Fit Statistic	Model	Vehicle Type					
		Auto	PUV	SUV	Trucks	Buses	Motorcycles
Mean Absolute Error (MAE)	Proposed	0.0482	0.0398	0.0133	0.0163	0.0016	0.0019
	Default	0.0720	0.0475	0.0150	0.0350	0.0018	0.0019
	% Higher error in default over proposed	49.38	19.35	12.78	114.72	12.50	0.00
Mean Percentage Absolute Error (MPAE)	Proposed	7.60	17.36	38.45	44.43	-	-
	Default	11.35	20.70	47.69	67.96	-	-
	% Higher error in default over proposed	49.34	19.24	24.03	52.96	-	-
Pseudo- R^2	Proposed	0.43					
	Default	0.03					

The MAE and MPAE results in Table 4 clearly indicate the superiority of the proposed model over the default model. The MPAE for SUVs and trucks are high even for the proposed model, but this is an artifact of very low VMT fractions of these two vehicle types on many links. The more significant observation is that there is a large improvement in

the fit of the proposed model relative to the default model in these two vehicle classes, especially for trucks. More generally, the percentage higher error in the default model compared to the proposed model is quite substantial across all vehicle types.

The pseudo- R^2 measure of the proposed and default models in Table 4 is another indicator of the superior performance of the proposed model. The results indicate that the default model is able to explain only 3 percent of the variation of link VMT fractions, while the proposed model is able to explain 44 percent of this variation. This result implies that roadway classification alone does not contribute much to explaining VMT mix, and that there are several other very important link and land-use attributes that should be considered in VMT mix analysis. This implication is, of course, also quite apparent from the results in Table 2.

7.2. Emissions Performance Evaluation

To evaluate the benefit of the proposed VMT mix model for emissions prediction, we applied the conversion factors developed in Section 6 to convert the actual and predicted count-based VMT mix fractions into VMT fractions defined by the EPA typology. Since the total volumes and the lengths of each link in the sample were known, we computed the actual and predicted link VMT accrued by each of the EPA vehicle classes. Finally, we applied areawide emission factors that varied by pollutant type and EPA vehicle class (obtained from the MOBILE5 model by the North Central Texas Council of Government Staff for the Dallas-Fort Worth area) to calculate the “actual” and predicted emissions for each pollutant type on each link. The evaluation of the ability of the proposed model and the “default” model to replicate “actual” link emissions was based on the same three criteria used for VMT

mix analysis, namely, mean absolute error (MAE), mean percentage absolute error (MPAE), and pseudo-R².

Table 5 provides the evaluation results. The MAE results in the table indicate the average link-level error in the emissions predictions (in grams), while the MPAE provides the average link-level percentage error in the emissions predictions. The results again indicate that the proposed model fits the data much better than does the default model. The MAE in the default model is about 24 to 30 percent higher than that obtained from the proposed model, while the MPAE in the default model is between 50 to 120 percent higher than that obtained from the proposed model, depending on the pollutant type (the MPAE in Table 5 for emissions is of an order lower than that in Table 4 for VMT mix because the magnitude of link emissions is very high compared to the VMT mix fractions; thus, a 1 percent error in emissions implies a much larger absolute error compared to a 1 percent error in VMT mix fractions).

Table 5: Emissions Performance Evaluation

Fit Statistic	Model	Pollutant Type		
		Carbon Monoxide (CO)	Volatile Organic Compounds (VOCs)	Oxides of Nitrogen (NOx)
Mean Absolute Error (MAE) in grams	Proposed	9992	684	6098
	Default	12754	853	8030
	% Higher error in default over proposed	27.64	24.76	31.70
Mean Percentage Absolute Error (MPAE)	Proposed	1.84	1.30	8.71
	Default	2.93	2.00	19.06
	% Higher error in default over proposed	59.24	53.40	118.82
Pseudo-R ²	Proposed	0.515		
	Default	0.004		

The pseudo- R^2 measure from the two models again emphasizes the superior performance of the proposed model. In summary, the improvement in VMT mix predictions by the proposed model does indeed translate to improved emissions estimation.

8. CONCLUSIONS

VMT mix, or the distribution of vehicles by weight and fuel type, is an important traffic-related parameter used in determining the composite mobile-source emissions on links of a network. The emissions factors (in grams per mile of vehicle travel for each pollutant) vary quite widely among different vehicle classes, and therefore the emissions analysis is very sensitive to the VMT mix. Consequently, it is important to develop methods that provide accurate VMT mix values at the individual link level.

Current approaches to VMT mix determination apply aggregate-level values across all links in the road network based on national-level traffic count statistics, or they apply roadway class-specific values based on local vehicle classification traffic counts. However, it has been documented in the literature that there is substantial variation in VMT mix across different regions and across links of the same roadway class within a region.

This paper proposes a fractional split model that relates VMT mix on links to several variables more informative and explanatory than merely roadway class. The fractional split model is a valuable formulation for VMT mix analysis because (1) it accommodates boundary values of the fractional VMT in a vehicle class; (2) it is easy to estimate using commonly available econometric software; and (3) it is easy to apply in forecasting mode to predict the VMT mix on each link of a network. A quasi-likelihood approach that provides

consistent and asymptotically robust inference for the parameters in the fractional split model is used in estimation.

The empirical analysis in the paper applies the fraction split model structure to estimate a VMT mix model for the Dallas-Fort Worth MA in Texas. Several data sources are used to assemble the data needed in the estimation. This assembly requires a reasonable, though not very substantial, amount of effort. Once the data are assembled, estimation of the VMT mix model proposed here is straightforward, as is the application of the model to predict link VMT mix values. Thus, though the current paper uses the Dallas-Fort Worth region as the study area, similar models can be estimated easily in other areas. This is particularly the case today because many metropolitan areas now have network and land-use files in GIS format, from which the information required for the proposed model estimation can be immediately extracted.

The empirical results for the Dallas-Fort Worth area show important differences in VMT mix based on link functional classification, link physical attributes, link speed characteristics, the degree of urbanization of the zone that contains the link, and land-use variables of the zone in which the link lies. Model evaluation efforts indicate that the proposed model provides much better predictions of VMT mix and emissions estimation than does the default model in use by Metropolitan Planning Organizations. The proposed model is currently being embedded within a GIS platform to predict the VMT mix on all links of the Dallas-Fort Worth MA.

There are two limitations in the current empirical analysis. First, variations in VMT mix across different times of day are not captured in the model. Second, seasonal variations in VMT mix are not incorporated in our model. The vehicle-classification counts used in the

current report provided only 24-hour counts and were conducted during the same month each year. Thus, they are inadequate for capturing temporal and seasonal variations. To accommodate these variations, more extensive vehicle-classification counts made at different times of day and different seasons of the year are needed. Once such data become available, the fractional split model structure can be applied to capture these additional effects. It is likely that the measures of fit of the proposed model in Tables 4 and 5 will improve even more after accommodating such temporal and seasonal variations in VMT mix.

ACKNOWLEDGMENTS

This study was funded by a TxDOT project titled “Transportation Control Measure Effectiveness Evaluation in Ozone Nonattainment Areas.” The authors would like to thank Bill Knowles, Carol Nixon, Wayne Young, and George Reeves for their valuable input throughout this research effort. Arnold Breedan of TxDOT’s Transportation Planning and Programming Division provided the vehicle-classification count data. Ken Cervenka, Mahmoud Ahmadi, and Gustavo Baez of NCTCOG provided and clarified data issues related to the road network file, the zonal land-use characteristics file, and the zonal coverage file for the Dallas-Fort Worth area. Ken Kirkpatrick and Christopher Klaus helped clarify current NCTCOG transportation air-quality procedures. Finally, the authors would also like to thank five anonymous referees for comments that considerably improved the presentation and technical aspects of this report.

REFERENCES

1. Chatterjee, A., Miller, T. L., Philpot, J. W., Wholley, Jr., T. F., Guensler, R., Hartgen, D., Margiotta, R.A., and P.R. Stopher, “Improving Transportation Data for Mobile-Source Emission Estimates,” National Cooperative Highway Research Program Report 394, Transportation Research Board, 1997.
2. *National Cooperative Highway Research Program Research Results Digest*, “Development of an Improved Framework for the Analysis of Air Quality and Other Benefits and Costs of Transportation Control Measures,” Transportation Research Board, National Research Council, March 1998.
3. Bhat, C. R., and R. Misra. “Discretionary Activity Time Allocation of Individuals between In-Home and Out-of-Home and between Weekdays and Weekends,” *Transportation*, Vol. 26, No.2, 1999, pp. 193–209.
4. Papke, L. E., and J. M. Wooldridge. “Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates,” *Journal of Applied Econometrics*, Vol. 11, 1996, pp. 619–632.
5. Maddala, G. S. *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge, 1983.
6. Bhat, C. R. “Modeling the Commute Activity-Travel Pattern of Workers: Formulation and Empirical Analysis,” Technical Paper, Department of Civil Engineering, The University of Texas at Austin, 2000.
7. Gourieroux, C., A. Monfort, and A. Trognon. “Pseudo-Maximum Likelihood Methods: Theory,” *Econometrica*, Vol. 52, 1984, pp. 681–700.

8. Wooldridge, J. M. "Specification Testing and Quasi-Maximum Likelihood Estimation," *Journal of Econometrics*, 48, 1991, pp. 29–55.
9. Ben-Akiva, M., and S. R. Lerman. *Discrete Choice Analysis*, MIT Press, Cambridge, 1985.
10. Davis, S. C. *Transportation Energy Data Handbook*, prepared by the Oak Ridge National Laboratory for the Office of Transportation Technologies, U.S. Department of Energy, Edition 17, 1997.