Household Planning of Transportation Use

Shuyi Jiang
Department of Management & Economics
Emmanuel College
United States of America
E-mail: jiangs@emmanuel.edu

Abstract

In this paper, I examine the household demand for travel by mode of transportation. I use a mixed logit model model to incorporate heterogeneity considerations. I estimate the mixed logit model by introducing two random coefficients based on a maximum simulated likelihood function. The demand framework used in this study is a class of differentiated product demand models, which describe bundles of product characteristics. I estimate travel demand, taking into account the demand for activity participation and interactions among household members using a differentiated product demand model. In this framework, individuals choose the travel alternatives to maximize their utility derived from the socio-demographic characteristics and the attributes of each alternative, following Berry, Levinsohn, and Pakes (1995). Empirical results demonstrate that individual and household socio-demographics are important and strong factors affecting individual choice of transportation. Work-related trip increases the use of private vehicle by 5.12%. School-related trip decreases the use of private vehicle by 3.35%. Senior citizens use private vehicle more often than public transportation. They increase the private vehicle use by 6.25%.

JEL classification: C25; C41; L11; L71; Q4

Keywords: transportation, demand, trip, multinomial logit, mixed multinomial logit

1 Introduction

In most metropolitan areas the increasing number of automobiles creates negative externalities such as air pollution, noise, depletion of energy reserves, and traffic congestion. Given the persistence of negative externalities related to automobile use, reducing the demand for car usage, especially in urban areas, is an important policy goal. In order to achieve this goal, the households need to make their car use more efficient and planners need to know whether households would like to switch transit modes from private vehicle to public transportation. Therefore, researchers need to consider how individual behavior affects transportation mode choices. A well-specified travel demand model is useful because it can be used to answer policy-related questions. For example, policy-makers would like to be able to address the following questions: What effects do personal income and car ownership have on travel demand and future congestion on each travel mode? What effects do travel time and cost changes have on total travel demand and on the demand for the different travel modes? Questions like these are the typical regional-level policy questions that planners would like to have answered, so it is necessary to have a model of how the household's socio-demographics and purpose of the trip affect the means of transportation used by the household.

The purpose of this study is to find out how household characteristics, and the purpose of day trips (including working, shopping, school trips or recreation) affect the chosen means of transportation, including private vehicle or public transportation. In this paper, I examine the household demand for travel by mode of transportation. In addition to examining a more recent time period than previous studies, an innovation of this paper is to include two policy variables to capture information about travel cost in the travel demand model. In addition, I use a mixed logit model to incorporate heterogeneity considerations. Mixed logit models have been widely adapted for this purpose (McFadden and Train, 2000). I estimate the mixed logit model by introducing two random coefficients based on a maximum simulated likelihood function. I identify the factors that determine household driving behavior by starting with a multinomial logit model. I use this as a benchmark for a comparison with a mixed logit model. The empirical results demonstrate that individual and household socio-demographics are important and strong factors affecting individual transportation choices. The remainder of the paper is as follows: Section 2 provides a brief discussion of travel demand models and the methodology. Section 3 discusses the survey data. Section 4 presents the multinomial logit model and mixed multinomial logit model specifications. Section 5 presents estimation results. Finally, a brief summary of the principal findings and conclusions follows in section 6.

2 Methodology and Travel Demand Models

This paper looks into the relationship between activity participation, such as work, shopping, school and recreation, and travel patterns, such as public transit and private vehicle use for all household members.

It develops a structural discrete choice model that will give distinguishable probabilities of daily pattern choices for each individual. The demand framework used in this study is a class of differentiated product demand models, which describe bundles of product characteristics. In this framework, individuals choose the travel alternatives that maximize their utility conditional on the socio-demographic characteristics and attributes of each alternative, following Berry, Levinsohn, and Pakes (1995, 2004). The foundations of these travel demand models were developed by McFadden (1974). Logit models are commonly used for forecasting the demand for alternatives, because the formula for logit choice probabilities is readily interpretable, especially compared to other types of models. In addition, the parameters of logit models are relatively easy to estimate. Behavioral travel-demand models are summarized in Domencich and McFadden (1975) and Stopher and Meyburg (1975). They review the developments in model specification, estimation, model evaluation and testing, and aggregation and forecasting. Models or hypotheses are formed on the nature of decision processes, and are evaluated in light of observed behavior. Domencich and McFadden (1975) outline a general procedure for formulating econometric models of population choice behavior from distributions of individual decision rules. They found that as auto time and cost increase relative to transit, the probability of choosing auto goes down. Respondent's race and occupation are important factors that affect probability of choosing auto for work.

Mixed logit models have been widely used to analyze consumers' travel-mode choice behavior in this research field (McFadden and Train, 2000). Among most of the available data analyzing transportation mode choices behavior, only one choice of one individual is recorded. Theoretically, with this type of cross-sectional data, it is not possible to distinguish between randomness and taste variations. As a result, it is not possible to estimate individual-specific parameters. McFadden and Train (2000) propose that it is feasible to capture the differences among people by assuming that the model parameters follow some distribution, instead of a fixed parameter as constrained in the closed form generalized extreme value (GEV) models, such as the multinomial logit and nested logit models which have been widely used to describe choice behavior in a variety of domains during the last twenty to thirty years. The analyst chooses the distributions and the parameters are obtained by the estimation procedure. The most commonly assumed distributions are normal, lognormal, and uniform distributions.

Individuals must decide which means of transportation to take, so any decision will fall into one of these three categories: private vehicle, public transit or other means of transportation. Thus, these three alternatives are the dependent variables in this analysis. The independent variables must either be exogenous to the estimation system such as vehicle attributes or be an output such as household socioeconomic characteristics. The independent variables in this study are trip distance, household total income per capita, age, gender, and race, vehicle per driver, travel cost and time per mile, and trip purposes. Trip purposes are included in this analysis, because individuals may choose different modes of transportation based on purpose of the trip . I aggregate the trip purposes into four categories: work-related trip, shopping trip, school trip and recreation trip.

3 The Survey Data and Variables

The survey dataset used to estimate the model is the 2001 National Household Travel Survey (NHTS). The survey provides information to assist transportation planners and policy makers. There are 69,817 usable households in the national sample. Travel days were assigned for all seven days of the week, including all holidays. The first travel day assigned was March 29, 2001. The last travel day assigned was May 4, 2002. The designated 24-hour travel day starts at 4:00 a.m. of the day assigned and continues until 3:59 p.m. of the following day. The NHTS is composed of different levels of data for (1) household level, which includes household size, age and gender of the household member, worker status of each household member, number of vehicles and income; (2) person level, which includes education level, worker status, and driver status; (3) vehicle level, vehicle model, model year, months vehicle owned, annual miles driven, and primary driver; (4) day trip level, which includes trip purpose, distance to destination, mode of transportation, vehicle used, and time trip began; (5) long distance trip level, which includes trip purpose, access and egress modes, overnight stops, transportation mode and stop purpose. I only focus on day travel dataset because my primary goal is to find how the household's socio-demographics and purpose of the day trip affect the means of transportation used by the household on a daily basis.

The variables included in my analysis will be the chosen mode of the trip, trip origin and destination, travel cost and time, and the socioeconomic characteristics of travelers and their households. I generate three discrete choices, which can be one of the following: (1) private vehicle, includes car, van, SUV, motorcycle, pickup truck, other truck, and RV; (2) public transportation, includes bus, train, subway, trolley, shuttle; (3) other type of transportation, includes walk, bicycle, sailboat, motorboat, yacht, ship, cruise, taxicab, limousine and others. Summary statistics for the variables used for this empirical analysis are shown in Table 1.

Trip distance, *Trip Miles*, is measured in miles. Its mean is 8.48 miles, ranging from 0 mile to 1200 miles. Household total income per capita, *Income per Capita*, is an index variable, ranging from 1 to 18. The increment of the index is \$5000. Thus, if *Income per Capita* is equal to 1, household total income per capita is less or equal to \$5000. If *Income per Capita* is equal to 2, household total income per capita ranges from \$5000 to \$9999. If *Income per Capita* is equal to 3, household total income per capita ranges from \$10,000 to \$14,999. If *Income per Capita* is equal to 18, household total income is more than or equal to \$100,000. The mean of this variable is 6.01. Respondent's gender, *Gender*, is a dummy variable, 0 for female and 1 for male. Respondent's age, *Age*, ranges from 0 to 88. Its mean is 39 years old. Household members are included. Vehicle per driver, *VPD*, is the number of vehicles divided by the number of drivers in each household. It ranges from 0 to 10.

Price per Mile is the travel cost of the trip per one mile. Travel cost is calculated as the money spent for the individual trip. For trips that use a private vehicle, travel cost is equal to the trip distance divided by respondent's vehicle's gas mileage, miles per gallon (MPG), and then multiply by the price of the fuel in respondent's state. For trips that use public transportation, travel cost is the transit fares of the trip. I first look into where the trip takes place by which public transportation and then find the average transit fare in that location. For other means of transportation, travel costs for walking and bicycling are assumed to be zero; travel costs for sailboat, motorboat, yacht, ship are equal to the trip distance divided by respondent's boat's diesel mileage, and then multiply by the price of the fuel in respondent's state; travel cost for cruise is the average cruise fare for that trip; travel costs for taxicab, limousine are the average price of the service per mile multiplies by the trip distance.

Time per Mile is the time spent on the trip per one mile. Its mean is 3.05 minutes per mile, ranging from 0 minutes per mile to 57 minutes per mile. Trip purposes, *Work*, *School*, *Shopping* and *Recreation*, are dummy variables. Four race groups are included, *White*, *Black*, *Asian* and *Hispanic*, also dummy variables. Other race groups include American Indian, Alaskan Native, Native Hawaiian, other Pacific Islander.

4 Multinomial Logit Model Specification

The logit model is the earliest and most widely used discrete choice model. Its popularity is due to the fact that the formula for the choice probabilities takes a closed form and is readily interpretable. By far most frequently-used specification is the multinomial logit model despite the fact that a potentially important drawback is the independence from irrelevant alternatives property (IIA). The IIA is a necessary and sufficient property of the multinomial logit model. It states that the ratio of the probabilities of choosing any two alternatives is independent of the attributes or the availability of a third alternative. I choose the multinomial logit model because one of the main independent variables is household total income per capita. Income is invariant across alternatives. For alternative-varying regressors the conditional logit model should be used, but when instead the regressors do not vary over alternatives, the multinomial logit model is used. The multinomial logit model specifies:

$$p_{ij} = \frac{e^{x_i'\beta_j}}{\sum_{k=1}^{K} e^{x_i'\beta_k}}, \quad j = 1,\ldots,m. \tag{1}$$

Where p_{ij} is the probability of individual i chooses alternative j. $\beta_1, \ldots, \beta_m$ are m vectors of unknown regression parameters.

There are m alternatives and the dependent variable y is defined to take value j if the jth alternative is taken, $j = 1, \ldots, m$. Define the probability that alternative j is chosen as

$$p_j = \Pr[y = j], \quad j = 1,\ldots,m. \tag{2}$$

Thus y_j equals one if alternative j is the observed outcome and the remaining y_k equals zero, so for each observation on y exactly one of y_1, y_2, ..., y_m will be nonzero. The multinomial density for one observation can then be written as

$$f(y) = p_1^{y_1} \times p_2^{y_2} \times \ldots \times p_m^{y_m} = \prod_{j=1}^{m} p_j^{y_j}. \tag{3}$$

where $p_j^{y_j}$ is the probability that alternative j is chosen.

The likelihood function for a sample of N independent observations is then

$$L_N = \sum_{i=1}^{N}\sum_{j=1}^{m} y_{ij}\, p_{ij}. \tag{4}$$

where p_{ij} is a function of parameters and regressors. Maximum likelihood estimation (MLE) is then used to obtain coefficient estimates $\acute{\epsilon}$ s.

Although most studies report only the regression coefficients of the model, instead of the more informative marginal effects, I will focus on marginal effects on the choice probabilities of a change in the regressor for a given individual. One obtains marginal effects by calculating

$$\frac{\partial p_i}{\partial x_{ik}} = \frac{e^{x_i'\beta}}{(1+e^{x_i'\beta})^2}\beta_k. \tag{5}$$

The dependent variable has three alternatives: vehicle, public transportation, and other type of transportation. The observed mode shares for the three modes considered are shown in Table 2. The parameter estimates for the multinomial logit model are presented in Table 3. Marginal effects are calculated to better interpret the multinomial logit model. Table 4 shows the marginal effects for private vehicle, public transportation and other means of transportation.

From Table 3, *Trip Miles* is statistically significant at the 1% level. From Table 4, it affects the probability of taking a private vehicle, or public transportation positively and affects the probability of taking other means of transportation negatively. If the trip distance increases by one mile, the probability of a respondent taking a private vehicle increases by 0.54%, the probability of taking a public transportation increases by 0.03% and the probability of taking other forms of transportation decreases by 0.03%.

Income per Capita is statistically significant at the 1% level. The increment of this variable is $5000, so when household total income per capita increases by $5000, it will increase the probability of taking a private vehicle by 0.07%, and it will decrease the probability of taking a public transit by 0.04% and decrease the probability of taking other types of transportation by 0.03%. Therefore, as household total income per capita goes higher, people are more likely to choose a private vehicle as their means of transportation.

Gender is statistically significant at the 1% level for private vehicle and significant at the 10% level for public transportation. It affects the likelihood of choosing a private vehicle negatively but a public transportation positively, which implies that man are less likely to take a private vehicle but more likely to take a public transportation to go for daily trips, such as work and work-related trips.

Age ranges from 1 to 88. Its parameter estimates are statistically significant at the 1% level. This variable affects the likelihood of taking a private vehicle positively and the likelihood of taking a public transportation negatively, which indicates that people are more reliant on private vehicles when they get older. When a person uses his vehicle habitually, it is difficult to change his behavior. I calculated the marginal effects from the multinomial logit model for elderly people whose age is from 61 to 88. The results are presented in Table

5. When age falls into the category of 61 to 88, one more year older increases individual's probability of choosing a private vehicle by 6.25% and increases his or her probability of choosing a public transit by 0.01%. Thus, elderly people are more reliant on the private vehicle than young adults. Vehicle per driver, *VPD*, is a measure of vehicle availability in each household. It is reasonable to assume that individuals' decisions on the means of transportation and vehicle ownership are related. The decision made to purchase a vehicle implies that an automobile in one's household is more desirable as one's means of transportation. The parameter estimates of *VPD* are statistically significant at the 1% level. *VPD* affects the likelihood of choosing a private vehicle positively and affects the likelihood of choosing a public transportation negatively. One more vehicles per driver in one's household, the probability of taking a private vehicle to go for a day trip increases by 10.22%, but the probability of choosing a public transportation decreases by 8.45%.

The parameter estimates of *Price per Mile* are significant at the 5% level. *Price per Mile* affects both the possibility of taking a private vehicle and public transportation negatively. If *Price per Mile* increases by 1 dollar/mile, it decreases the likelihood of choosing a private vehicle by 3.22% and decreases the likelihood of choosing a public transportation by 6.31%.The parameter estimates of *Time per Mile* are significant at the 10% level. If *Time per Mile* increases by 1 minute/mile, it decreases the likelihood of taking a private vehicle by 9.05% but increases the likelihood of taking a public transportation by 7.10%. I aggregate the purpose of the day trips into five categories: work trip, school trip, shopping trip, recreation trip and other types of trips. Each of these four variables are statistically significant at the 1% level relative to other types of trips.

Work Trip affects the likelihood of taking private vehicle and public transit positively. It increases the probability of the private vehicle by 5.12% but only increases the probability of taking a public transit by 0.25%.

School Trip affects the probability of taking a private vehicle negatively by 11.63% and the probability of taking a public transportation positively by 11.73%. The signs are expected, because most children go to school by the school bus.

Shopping Trip has the expected signs as well. It affects the probability of taking a private vehicle positively by 3.35% and the probability of taking a public transportation negatively by 2.73%. Shoppers find taking public transportation is more onerous than taking their own vehicle because they may have quite a number of bags after shopping.

Recreation Trip affects the probability of taking a private vehicle negatively by 9.48% and the probability of taking public transportation negatively by 1.14%, which implies that people are more likely to take other types of transportation instead of their own vehicle or public transportation for daily trips. Since there are high frequencies of transportation modes that are walking, bicycling and boating, people may choose these alternatives for recreation purposes within one day. For ethnic backgrounds, I include *White*, *Black*, *Asian* and *Hispanic* in the estimation. *White, Black* and

Hispanic are statistically significant at the 1% level. *Asian* is not significant. The problem with the multinomial logit model is that all the parameters are assumed to be fixed. If the non-intercept coefficients of the independent variables depend on unobserved features of the trip-maker, then random coefficient variation arises in the travel mode choice model. Therefore, mixed multinomial logit model is introduced in the following section.

5 Mixed Multinomial Logit Model Specification

Consider a mixed multinomial logit model that allows correlation and heteroscedasticity (McFadden and Train, 2000). The model assumes that a person faces a choice of j alternatives. His or her utility from any alternative can be decomposed into a non-stochastic part, linear in parameters, that depends on observed data, a stochastic part that is correlated over alternatives and heteroscedastic, and the last part that is independently identically distributed (i.i.d.) over alternatives and people. I follow the traditional discrete choice random coefficients literature that includes Domencich and McFadden (1975) and Hausman and Wise (1978). This literature assumes a linear utility function u_{ij} obtained by consumer i from the choice of product j. When $j = 0$, u_{i0} is the utility of the consumer if he does not purchase any of the J goods and allocates all income to other goods. The utility of individual i choosing mode j for trip t is

$$U_{ijt} = X_{ijt}\beta_{ijt} + \eta_{jt} + \varepsilon_{ijt}. \tag{6}$$

where x_{ijk} is a vector of observed variables related to alternative j and the consumer i in trip t.

Further, let k denote the observed product characteristics. The model then becomes

$$u_{ijt} = \sum_k x_{ijt}\beta_{ik} + \eta_{jt} + \varepsilon_{ijt}. \tag{7}$$

β_{ik} is a vector of parameters that are fixed over consumers and alternatives. η_{jt} is a random error term with mean zero, whose distribution over consumers and alternatives depends on parameters related to alternative j in trip t. ε_{ijt} is a random error term with mean zero and is i.i.d. over alternatives.

Assuming each ε_{j} is independently, identically distributed extreme value, the logit formula can be expressed as:

$$L_i(\eta) = \frac{e^{x_i^{'}\beta}}{\sum_j e^{x_j^{'}\beta + \eta_j}}. \tag{8}$$

With the mixed multinomial logit model, the choice probability is a mixture of logits, with density function η of as the mixing distribution. Since η is not given, the unconditional choice probability is the above logit formula integrated over all values of η weighted by the density of η:

$$P_i = \int L_i(\eta) f(\eta \mid \Omega) d\eta. \tag{9}$$

The probabilities do not exhibit the IIA property. Here $f(\beta|\theta)$ is the density of β, where θ represents the random parameters of the distribution for β. Define $\eta_j = \beta z_j$. Since η_j is the unobserved random portion of utility, it can be correlated over alternatives depending on the specification of z_j. Researchers specify a distribution for coefficients and estimate the parameters of that distribution. The random coefficient β of variable z_j represents the variation over people. For the standard multinomial logit model, z_j is identically equal to zero. Thus there is no correlation over alternatives in the utility function. This lack of correlation gives rise to the IIA property. With nonzero random coefficients, utility is correlated over alternatives, which introduces a heteroscedasticity problem.

Since the formula (9) does not have a closed form, it is not possible to calculate it directly, requiring simulation methods to approximate it. Given a value for parameter θ, a value of β is drawn from its distribution. Thus the logit solution can be obtained by simulation. This process is repeated for many times, at least 100 draws roughly. So, the simulated probability that the person chooses alternative i will be

$$SP_i = \frac{1}{R}\sum_{r=1,\dots,R} L_i(\eta^r). \tag{10}$$

where R is the number of draws to repeat. The simulated probabilities sum to one over alternatives, which is useful in forecasting the travel demand model.

A mixed multinomial logit is well suited to simulation methods for estimation. If the simulated probabilities are inserted into the log-likelihood function, the simulated log-likelihood is obtained:

$$SLL = \sum_{n=1}^{N}\sum_{j=1}^{J} d_{nj} \ln SP_i. \tag{11}$$

where $d_{nj} = 1$ if individual n chooses alternative j and $d_{nj} = 0$ otherwise. The maximum simulated likelihood estimator is the value of θ that maximizes SLL.

Parameter estimates for the mixed multinomial logit model are presented in Table 6. The multinomial logit (MNL) model is estimated using maximum likelihood method and the mixed multinomial logit (MXL) model is estimated using maximum simulated likelihood with a normal distribution assumption on the two parameters of *VPD*. The maximum simulated likelihood is obtained using the GAUSS programming language. The MNL model provides only point estimates for each of these two parameters. The standard errors for these point estimates are low. The parameter estimates and the standard errors are somewhat different from the ones obtained by the mixed multinomial logit model. With 100 repetitions, the estimated results for the MXL model give a standard deviation of variable *VPD* for private vehicle of 3.10, which is statistically significant at the 1% level, and for public transportation is 0.13, which is not significant. These results indicate that the heterogeneity assumption on *VPD* for a private vehicle is necessary but for public transportation may not be necessary.

The last part of Table 6 shows the measure of goodness-of-fit, including Log-Likelihood values at convergence and Bayesian Information Criteria (BIC) values. BIC is computed as a function of log-likelihood value at convergence, with a penalty on the number of parameters (Schwartz, 1978). If LL denotes the log-likelihood at convergence, K denotes the number of parameters estimated in the model and $NOBS$ denotes the total number of observation used to estimate the model, then $BIC = -2 \cdot LL + \ln(NOBS) \cdot k$. The model selection based on BIC, the model minimizing the BIC value should be selected. In terms of the comparison based on the goodness-of-fit measures, the Log-Likelihood value of the MXL model is better than for the MNL model, and the BIC value favors the MXL model, which means that heterogeneity problem needs to be considered in the parameters.

Two policy questions can be answered in this study: (1) What effects do personal income and car ownership have on travel demand and future congestion on each travel mode? The household income per capita has positive effects on the travel demand and demand of travel modes of a private vehicle and public transportation. (2) What effects do travel time and cost changes have on total travel demand and on the demand for the different travel modes? Compared to other means of transportation, travel time affects the likelihood of choosing a private vehicle negatively and public transit positively. Travel cost affects the likelihood of taking a private vehicle and public transportation negatively.

6 Conclusions

Metropolitan life depends on a healthy and efficient transportation system. Today's travel networks are automobile-centered, which leads to traffic, air pollution, and energy depletion. It is not possible to build enough roads to meet travel demand, even if the economic, environmental and health costs were ignored. Aware of these problems, reducing the demand for private vehicles is one solution. This analysis studies the choice of travel mode to understand individual behavior. I use a multinomial logit model to estimate how individual behavior influences the choice of transportation. I impose a restriction that constrains the coefficients of trip miles to be equal for both private vehicles and public transit. Using a likelihood ratio test, I find the restriction holds at the 1% level, which makes my model becomes the restricted MLE. In order to interpret the coefficients, I calculate the marginal effects from the efficient restricted model.

The empirical results show that individual and household socio-demographics are important and strong factors affecting individual choice of transportation. The need for motorized travel, or to drive longer distances, should be reduced by some other choices like walking, cycling and public transportation. From the results of marginal effects, the elderly people are more likely to take a private vehicle. One possible reason is habitual car usage. It has frequently been noted that daily travel patterns tend to repeat themselves from day to day, from week to week, and perhaps from year to year (Pendyala et al. 2000). Another possible reason is that elderly people are not aware of transportation alternatives. There are some special public transit services for seniors in most of communities. Kostyniuk and Shope (1999) show that half of the respondents are not aware of public or charitable options when asked if there would be other forms of transportation available for seniors if they could not drive.

Individuals are more likely to take a private vehicle for work, work-related trips, and shopping trips. People are more likely to go to work by public transit if the company is located in a downtown area, simply because it is convenient and the parking and road pricing are costly. Phang and Toh (1997) found that the introduction of charges in Singapore reduced vehicle volumes in the city center by 45%. Unfortunately, there are few real-world examples of road pricing for travel-demand management. However, there is a larger body of research on the use of parking charges to manage demand and influence mode choice. Flannelly et al. (1991) and Kuppam et al. (1998) found that increased parking costs had a greater influence on mode choice than incentives to carpool or take a transit. Multinomial logit model does not allow heteroscedasticity. Its specifications make overly restrictive assumptions about heteroscedasticity. In order to incorporate heterogeneity considerations into the study of mode choice, a mixed multinomial logit model is performed in the above behavioral choice analysis using a maximum simulated likelihood estimation. Thus, a comparison of the two is necessary so as to identify the best model. I calculated the Bayesian Information Criteria value for both models. The results favor the mixed multinomial logit model, so heterogeneity should be considered in the parameters.

References

Berry, S, J. Levinsohn, and A. Pakes, 1995. Automobile Prices in Market Equilibrium, *Econometrica, 63*: 841-990.

Berry, S, J. Levinsohn, and A. Pakes, 2004. Differentiated Products Demand Systems froma Combination of Micro and Macro Data: The New Car Market, *Journal of Political Economy, 112(1)*: 68-105.

Domencich, T. and D. McFadden, *Urban Travel Demand: A Behavioral Analysis*, Amsterdam: North Holland, 1975.

Flannelly, K. J., M. S. Jr. McLeod, L. Flannelly, and R. W. Behnke, 1991. Direct Comparison of Commuters' Interests in Using Different Modes of Transportation. *Transportation Research Record, 1321*: 90-96.

Hausman, J. A. and D. Wise, 1978. A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preference, *Econometrica, 46*: 403-426.

Kostyniuk, L.P. and J.T. Shope, 1999. *Choice of Transportation Mode Among Older Drivers and Former Driver*s. UMTRI-99-45.

Kuppam, A.R., R.M. Pendyala, and M.A.V. Gollakoti, 1998. Stated Response Analysis of the Effectiveness of Parking Pricing Strategies for Transportation Control. *Transportation Research Record, 1649*: 39-46.

McFadden, D., 1974. The Measurement of Urban Travel Demand, *Journal of Public Economics, 3*: 303-328.

McFadden, D. and K. Train, 2000. Mixed MNL Models for Discrete Response, *Journal of Applied Econometrics, 15(5)*: 447-470.

Pendyala, R. M., V. N. Shankar, and R.G. McCullough, 2000. Freight Travel Demand Modeling: Synthesis of Approaches and Development of a Framework. *Journal of the Transportation Research Board, 1725*: 9-16.

Phang, S. and R. S. Toh, 1997. From manual to electronic road congestion pricing: The Singapore experience and experiment. *Journal of Transport Geography, 10(1)*: 29-38.

Stopher, Peter R. and A. H. Meyburg, *Behavioral Travel-Demand Models*, Lexington Books, 1975.

Schwarz, Gideon E., 1978. Estimating the dimension of a model. *Annals of Statistics, 6(2)*: 461--464.

Table 1: Summary Statistics for all variables

Variable	Definition	Mean	St. Dev.	Min.	Max.
Vehicle	Private vehicle	0.79	0.41	0	1
Public Transportation	Public transportation	0.06	0.23	0	1
Other Transportation	Other means of transportation	0.15	0.36	0	1
Trip Miles	Trip distance (mile)	8.48	20.20	0	1200
Income per capita	Household total income per capita (dollar)	6.01	4.29	1	18
Gender	Respondent's gender	0.47	0.50	0	1
Age	Respondent's age	39.11	20.82	0	88
VPD	Vehicle per driver	1.00	0.49	0	10
Price per Mile	Travel cost per mile (dollar per mile)	0.05	0.19	0	0.43
Time per Mile	Travel time per mile (minute per mile)	3.05	5.83	0	57
Work	Work and work related trip	0.17	0.37	0	1
School	School trip	0.05	0.21	0	1
Shopping	Shopping trip	0.13	0.33	0	1
Recreation	Recreation trip	0.11	0.31	0	1
White	Caucasian	0.77	0.42	0	1
Black	African American	0.10	0.30	0	1
Hispanic	Hispanic	0.04	0.20	0	1
Asian	Asian	0.03	0.17	0	1

Table 2: The Share of Travel Modes

Travel Mode	Mode Share			
	Work Trip	School Trip	Shopping Trip	Recreation Trip
Private Vehicle	87.2%	60.2%	85.1%	68.4%
Public Transportation	6.2%	23.9%	2.5%	3.5%
Other Means of Transportation	6.6%	15.8%	12.4%	28.0%

Table 3: Estimates of Multinomial Logit Model

Private Vehicle (PV) and Public Transportation (PT)		
Variables	Parameter Estimates	Standard Error
Trip Miles_PV	0.0423***	0.0057
Trip Miles_PT	0.0425***	0.0057
Income per Capita_PV	0.0066***	0.0020
Income per Capita_PT	0.0017***	0.0033
Gender_PV	-0.1075***	0.0202
Gender_PT	0.0248*	0.0325
Age_PV	0.0132***	0.0005
Age_PT	-0.0107***	0.0008
VPD_PV	1.6711***	0.0254
VPD_PT	-0.4329***	0.0407
Price per Mile_PV	-0.0715**	0.0119
Price per Mile_PT	-0.0633**	0.0251
Time per Mile_PV	-0.0029*	0.0018
Time per Mile_PT	0.0073*	0.0041
Work Trip_PV	1.0320***	0.0336
Work Trip_PT	0.9163***	0.0475
School Trip_PV	-0.0535***	0.0476
School Trip_PT	1.5122***	0.0534
Shopping Trip_PV	0.6370***	0.0324
Shopping Trip_PT	-0.3389***	0.0622
Recreation Trip_PV	-0.9111***	0.0288
Recreation Trip_PT	-1.2847***	0.0599
White_PV	0.1270**	0.0422
White_PT	-0.2836***	0.0631
Black_PV	-0.2469***	0.0493
Black_PT	0.2102***	0.0691
Hispanic_PV	-0.3496***	0.0595
Hispanic_PT	-0.2782***	0.0854
Asian_PV	0.0331	0.0694
Asian_PT	0.0680	0.1000
Constant_PV	-1.4006***	0.0537
Constant_PT	-1.2165***	0.0772
Log-Likelihood	-49028	
Bayesian Information Criteria	98402	

***, Significant at the 1% level.
**, Significant at the 5% level.
*, Significant at the 10% level

Table 4: Marginal Effects for Multinomial Logit Model

Marginal Effects	Private Vehicle	Public Transportation	Other Means of Transportation
Trip Miles	0.540%	0.033%	-0.574%
Household Total Income per Capita	0.065%	-0.038%	-0.026%
Gender	-0.725%	0.566%	0.159%
Age	0.085%	-0.075%	-0.010%
Vehicle per Driver	10.219%	-8.451%	-1.768%
Price per Mile	-3.222%	-6.314%	0.255%
Time per Mile	-9.050%	7.099%	-1.392%
Work Trip	5.117%	0.245%	-0.757%
School Trip	-11.631%	11.732%	-0.101%
Shopping Trip	3.347%	-2.730%	-0.617%
Recreation Trip	-9.475%	-1.144%	2.091%
White	1.914%	-1.763%	-0.151%
Black	-2.416%	2.083%	0.332%
Hispanic	-0.812%	0.261%	0.551%
Asian	-0.093%	0.139%	-0.046%

Table 5: Marginal Effects for Multinomial Logit Model when Age is from 61 to 88 years old

Marginal Effects	Private Vehicle	Public Transportation	Other Means of Transportation
Trip Miles	0.758%	0.017%	-0.774%
Household Total Income per Capita	0.054%	0.021%	-0.032%
Gender	0.264%	-0.200%	-0.064%
Age (61~88)	6.253%	0.008%	-0.030%
Vehicle per Driver	7.431%	-4.153%	-3.278%
Price per Mile	-4.579%	-7.853%	0.270%
Time per Mile	-6.005%	7.207%	-4.598%
Work Trip	1.146%	0.111%	-1.257%
School Trip	1.287%	-0.657%	-0.630%
Shopping Trip	0.963%	-0.296%	-0.667%
Recreation Trip	-4.915%	0.749%	4.166%
White	3.326%	-1.886%	-1.440%
Black	0.325%	0.372%	-0.697%
Hispanic	-2.563%	1.933%	0.630%
Asian	-0.831%	1.292%	-0.461%

Table 6: Estimates of Mixed Multinomial Logit Model

Private Vehicle (PV) and Public Transportation (PT)		
Variables	Parameter Estimates	Standard Error
Trip Miles_PV	0.0430***	0.0063
Trip Miles_PT	0.0436***	0.0064
Income per Capita_PV	0.0080***	0.0032
Income per Capita_PT	0.0031***	0.0034
Gender_PV	-0.1075***	0.0293
Gender_PT	0.0248*	0.0337
Age_PV	0.0132***	0.0008
Age_PT	-0.0107***	0.0008
VPD_PV	1.6711***	0.0298
VPD_PT	-0.4329***	0.0495
Price per Mile_PV	-0.0776***	0.0248
Price per Mile_PT	-0.0689***	0.0406
Time per Mile_PV	-0.0032*	0.0022
Time per Mile_PT	0.0077*	0.0043
Work Trip_PV	1.0320***	0.0354
Work Trip_PT	0.9163***	0.0498
School Trip_PV	-0.0535***	0.0513
School Trip_PT	1.5122***	0.0552
Shopping Trip_PV	0.6370***	0.0410
Shopping Trip_PT	-0.3389***	0.0673
Recreation Trip_PV	-0.9111***	0.0329
Recreation Trip_PT	-1.2847***	0.0606
White_PV	0.1264**	0.0424
White_PT	-0.2820***	0.0640
Black_PV	-0.2308***	0.0533
Black_PT	0.2349***	0.0748
Hispanic_PV	-0.3467***	0.0715
Hispanic_PT	-0.2663***	0.0932
Asian_PV	-0.0382	0.0866
Asian_PT	0.0685	0.1005
Constant_PV	-1.3315***	0.0541
Constant_PT	-0.4571***	0.0774
Standard Deviation for Vehicle per Driver		
VPD_ PV	3.0952***	0.0859
VPD_ PT	0.1341	0.0773
Log-Likelihood	-49548	
Bayesian Information Criteria	99466	

***, Significant at the 1% level.
**, Significant at the 5% level.
*, Significant at the 10% level.