The Delta Method to Compute Confidence Intervals of Predictions from Discrete Choice Model: An Application to Commute Mode Choice Model

Xin Ye, Ph.D.
(Corresponding Author)
Assistant Professor
Civil Engineering Department
California State Polytechnic University
3801 West Temple Avenue, Pomona, California 91768
Phone: 909-869-3444
Email: ye@csupomona.edu

Ram M. Pendyala, Ph.D.
Professor
School of Sustainable Engineering and the Built Environment
Ira A. Fulton Schools of Engineering
Arizona State University
PO Box 875306
Tempe, AZ 85287
Phone: 480-727-9164
Email: Ram.Pendyala@asu.edu

Submitted to the 2014 Annual Meeting of the Transportation Research Board for Presentation Only (Committee ADB40)
Word Count: 3,673 text + 4 tables x 250 + 4 figures x 250 = 5,673 words

Abstract

In this paper, we suggest using the Delta method to compute confidence intervals of predictions from discrete choice model. This method is applied to a commute mode choice model based on multinomial logit modeling method. The model is estimated and tested based on a sample of 3408 work trips collected from the Aargau Canton of Switzerland in 2000. About 70% of trips are randomly selected for model estimation and the rest 30% are used for testing the Delta method. Monte Carlo simulation is conducted to generate the confidence intervals of predictions from the estimated discrete choice model. It is found that the Delta method can replicate the confidence intervals obtained from simulations almost perfectly. This result demonstrates that the Delta method is a numerically accurate and computationally efficient approach to compute confidence intervals of predictions from a discrete choice model.

Key Words: Discrete Choice Model, Multinomial Logit Model, Confidence Interval, Delta Method, Travel Mode Choice Model, Choice Frequency Prediction

Background

The discrete choice modeling method based on individual's utility maximization principle was proposed in 1970' and was immediately applied to model travelers' choice behaviors (Mcfadden, 1973). Up to now, discrete choice modeling method has been applied to model almost every aspect of travel behavior and therefore plays a central role in travel demand models. In the state-of-practice trip-based travel demand model, destination choice and travel mode choice models have become classic examples for successful application of discrete choice models (Mcfadden, 1973; Ben-Akiva and Lerman, 1985; Bhat et al., 1998). Moreover, in the state-of-art activity-based travel demand model, discrete choice models have also been widely used for modeling travelers' activity behaviors such as activity type choice, activity timing, etc. (Gangrade et al., 2002; Ye and Pendyala, 2009).

Discrete choice models receive wide acclaim in transportation area but the modeling method is applied in different ways. In many situations, the discrete choice modeling method is regarded as a multivariate statistical tools for discrete data analysis. Modelers' main purpose to develop models is to understand how some explanatory variables affect choice decisions. Since the impacts of explanatory variables on choice variables are directly quantified by mode coefficients, modelers are more concerned about the signs, magnitudes, significances, errors of their model coefficients. For example, the ratio between two coefficients reflects the value of decision makers on two explanatory variables associated with those two coefficients. In some studies, sensitivity analysis is conducted based on estimated models to better illustrate the impacts of explanatory variables on choice probabilities (e.g. Pinjari et al., 2008).

However, in some occasions, modelers apply discrete choice modeling method as a predictive tool (e.g. Zhang and Xie, 2008). Discrete choice models can be applied at either individual (disaggregate) or population (aggregate) levels. It should be admitted that it is extremely difficult to accurately predict individual's travel choice behavior given the currently available data. However, accumulated individuals' choices, reflected by the market shares (e.g. travel mode shares), can be predicted at a certain accuracy level. A good example of such an application can be found in Train (pp. 74, 2009). When a discrete choice model is applied at individual level, it only generates a probability for a traveler to choose an alternative. Strictly speaking, the probability generated from a discrete choice model should not be considered as a prediction of traveler's choice.

The purpose of this paper is to present a method to apply a discrete choice model to conduct prediction in a rigorous and scientific way. Since the objects of travel behavior research are human beings, highly accurate predictions are usually not expected. The prediction from discrete choice model needs to be considered as a random variable, which follows a certain distribution and has expectation and variance. A confidence interval needs to be created to show the lower and upper bounds of the prediction from a discrete choice model. It is the primary objective of this paper.

The remainder of this paper is organized as follows. In the next section, relevant literature and methods will be reviewed. Then, in the section of methodology, the Delta method will be presented for computing confidence interval of predictions from a discrete choice model. In the section of case study, a commute mode choice model is estimated based on 2000 Swiss Microcensus travel survey data to demonstrate the applicability and correctness of the Delta method. Conclusions and discussions will be made finally.

Literature Review

This section consists of two parts. In this first part, we will review the method to compute confidence interval of prediction from a classical linear regression model. In this regard, discrete choice model is analogous to a linear regression model. The review can aid in a better understanding of the mechanism to derive predication's variance and confidence interval. In the second part, we will review earlier work in transportation literature that has made effort to solve similar problems.

1. Review of Prediction's Variance and Confidence Interval from Linear Regression Models

Linear regression model starts from the assumption that $Y = X\beta + \varepsilon$, where X is a n × k matrix of observed independent variables (it may contain a vector of one for a constant term) and Y is a n × 1 vector of observed dependent variable; ε is a random disturbance to capture unobserved or unspecified random variables whose expectation is 0 and variance is σ^2; β is a k × 1 vector of coefficients associated with explanatory variables. "n" is the total number of observations and "k" is the total number of coefficients to be estimated. The classical least square estimation method will provide unbiased and consistent estimators $\hat{\beta} = (X'X)^{-1}XY$ when $E(\varepsilon|X) = 0$. Suppose there is a 1 × k vector of observed x_1 outside the sample X for model estimation, the prediction of dependent variable can be represented as $\hat{y}_1 = x_1\hat{\beta} + \varepsilon_1$. By taking expectation on both sides of the equality, $E(\hat{y}_1) = E(x_1\hat{\beta}) + E(\varepsilon_1) = x_1E(\hat{\beta})$. Since $\hat{\beta}$ is unbiased estimators of β, $E(\hat{y}_1) = x_1\beta$. The variance of $\hat{y}_1$ can be expressed as $\mathrm{Var}(\hat{y}_1) = \mathrm{Var}(x_1\hat{\beta}) + \mathrm{Var}(\varepsilon_1) = x_1\mathrm{Var}(\hat{\beta})\,x_1' + \hat{\sigma}^2$, where $\hat{\sigma}$ is an unbiased estimator of σ, calculated as $\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i\hat{\beta})^2}{n-k}}$. $\mathrm{Var}(\hat{\beta})$ is a variance-covariance matrix of $\hat{\beta}$ and $\mathrm{Var}(\hat{\beta}) = \hat{\sigma}^2(X'X)^{-1}$. Based on the assumption that ε follows a normal distribution, $\hat{y}_1$ also follows a normal distribution. Thus, when n >> k, (1 - α) × 100% confidence interval of $\hat{y}_1$ is $x_1\hat{\beta} \pm \Phi\left(1-\frac{\alpha}{2}\right)\cdot\sqrt{x_1\mathrm{Var}(\hat{\beta})x_1' + \hat{\sigma}^2}$, where $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution (Greene, 2002). When α is 0.05, $\Phi^{-1}\left(1-\frac{\alpha}{2}\right) \approx 1.96$. A similar method can be applied to generate confidence interval of predictions from a discrete choice model.

2. Review of Transportation Literature on Similar Topics

The earliest work in transportation literature on this topic is the one by Horowitz (1979). In this work, asymptotic sampling method, linear programming and non-linear programming methods are applied to develop confidence intervals for the choice probabilities from multinomial logit models. However, those methods have not been widely applied in transportation literature. Amstrong et. al. (2001) proposed two methods to solve a similar problem about how to obtain a confidence interval of value of time based on a choice model. Cherchi (2009) stressed the importance of confidence of prediction from discrete choice models. Daly et. al. (2013) proposed using the Delta method for calculating errors for measures derived from choice modeling estimates. The same Delta method can also be used to accurately estimate variance of a prediction from a discrete choice model and thereby aid in generating confidence interval of the prediction.

Methodology

In this section, multinomial logit model (MNL) will be used as an example of discrete choice model to present the Delta method but this method should be applicable to any kinds of discrete choice models. In MNL, the probability of individual "i" choosing alternative "j" can be expressed as: $P_{ij} = \frac{\exp(x_{ij}\beta_j)}{\sum_{k=1}^{J}\exp(x_{ik}\beta_k)}$, where linear-in-parameter utility functions are specified and x_{ij} is a vector of explanatory variables for alternative "j" and individual "i"; β_j is a vector of coefficients associated with variables x_{ij} for alternative "j"; "J" is the total number of alternatives in the choice set. In the model estimation process, Maximum Likelihood Estimation (MLE) method is applied to consistently and efficiently estimate all the coefficients in vectors β_j. If all the coefficients in vectors β_j are denoted as a vector β, the estimator of β is denoted as $\hat{\beta}$. According to the theorem of MLE, $\hat{\beta}$ is asymptotically normally distributed with the expectation of β and the variance-covariance matrix of the negative inverse of the Hessian matrix [i.e. -$H^{-1}(\hat{\beta})$].

Since a discrete choice model only provides a probability of each alternative being chosen but does not directly predict a choice of an individual, we only consider the predicted choice frequency of one alternative within a sample (i.e. accumulated individual's probability over a sample) in this paper. The prediction of choice frequency from a discrete choice model can be expressed as the sum of calculated probabilities over a sample:

$$F_j = \sum_{i=1}^{N} P_{ij} \ , \tag{1}$$

where "i" is an index for individual, "j" is an index for alternative, "F_j" represents the frequency of alternative "j" being chosen in a sample, "N" is the sample size.

If inputting the probability formula of MNL: $P_{ij} = \frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}$ into Equation 1, one may obtain that

$$F_j = \sum_{i=1}^{N}\left[\frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}\right]. \tag{2}$$

Since estimators in $\hat{\beta}$ are random variables, the predicted choice frequency F_j is also a random number. To calculate the confidence interval of F_j, one needs to know the distribution of F_j. Monte Carlo simulation method can always be used to draw random numbers for $\hat{\beta}$ and input into Equation (2) to calculate F_j. If this process is repeated many times, the distribution of F_j can be simulated and its confidence interval can be estimated. However, this process is very time-consuming. Time consumption may not be an extremely serious problem for a discrete choice model with closed-form probability formula but computational efficiency will be an issue for models with no closed-form probability formula (e.g. multinomial probit model, mixed logit model). Thus, a more computationally efficient method is desired for those complex choice models.

The Delta method, introduced in Greene (2002), can be applied to approximate the distribution of F_j, which is a function with respect to $\hat{\beta}$ following a multivariate normal distribution. That is because the Delta method can be applied to derive an approximate probability distribution for a function of estimators which are asymptotically normally distributed. The following is the theorem of the Delta method: If $(\hat{\beta}-\beta) \xrightarrow{d} N[0, Var(\hat{\beta})]$ and f($\hat{\beta}$) is a continuous function with respect to $\hat{\beta}$, then

$$f(\hat{\beta}) \xrightarrow{d} N\left[f(\beta), g(\hat{\beta})^T Var(\hat{\beta}) g(\hat{\beta})\right], \quad (3)$$

where $g(\hat{\beta})$ is the first-order partial derivative $\partial f(\hat{\beta})/\partial\hat{\beta}$, which is a column vector, and N(μ, Σ) represents a multivariate normal distribution associated with expectation μ and variance-covariance matrix Σ.

Here, we quickly show Formula (3) for better understanding the Delta method. One may first expand $f(\hat{\beta})$ into a Taylor series: $f(\hat{\beta}) = f(\beta) + \frac{f'(\beta)}{1!}(\hat{\beta}-\beta) + \frac{f''(\beta)}{2!}(\hat{\beta}-\beta)^2 + \cdots + \frac{f^{(k)}(\beta)}{k!}(\hat{\beta}-\beta)^k$. Since $\hat{\beta}$ is a consistent estimator of β, $\hat{\beta}$ should be very close to β if the sample size is large enough for model estimation. The items with high power terms should be negligible. If only taking the first two items, one can obtain $f(\hat{\beta}) \approx f(\beta) + f'(\beta)(\hat{\beta}-\beta)$. Since $\hat{\beta}$ follows a normal distribution and $f(\hat{\beta})$ is almost a linear transformation of $\hat{\beta}$, $f(\hat{\beta})$ follows a normal distribution as well. Using *g(β)* instead of $f'(\beta)$, we can rewrite the equation as $f(\hat{\beta}) = f(\beta) + g(\beta)(\hat{\beta}-\beta)$. The expectation $E[f(\hat{\beta})] = E[f(\beta)] + g(\beta)E(\hat{\beta}-\beta)$. $\hat{\beta}$ is an unbiased estimator of β, $E(\hat{\beta}) = \beta$ and then $E(\hat{\beta}-\beta) = 0$. Therefore, $E[f(\hat{\beta})] = f(\beta) \approx f(\hat{\beta})$.

The variance $Var[f(\hat{\beta})] = Var[f(\beta) + g(\beta)(\hat{\beta}-\beta)]$. Since β, $f(\beta)$ and $g(\beta)$ are constants, their variances are zero. The variance of $f(\hat{\beta})$ can be simplified as $Var[f(\hat{\beta})] = Var[g(\beta)\hat{\beta}] = g(\beta)^T Var(\hat{\beta}) g(\beta) \approx g(\hat{\beta})^T Var(\hat{\beta}) g(\hat{\beta})$, where $g(\hat{\beta})$ is the first-order partial derivative $\partial f(\hat{\beta})/\partial\hat{\beta}$. Thus far Formula (3) has been derived.

To compute the confidence interval of predicted choice frequency F_j based on a multinomial logit model, one may consider the function $f(\hat{\beta}) = \sum_{i=1}^{N}\left[\frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}\right]$ as per Equation (2). According to the Delta method, F_j approximately follows a normal distribution, $E(F_j) = \sum_{i=1}^{N}\left[\frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}\right]$ and $Var(F_j) = g(\hat{\beta})^T Var(\hat{\beta}) g(\hat{\beta})$, where $g(\hat{\beta}) = \frac{\partial f(\hat{\beta})}{\partial\hat{\beta}} = \frac{\partial \sum_{i=1}^{N}\left[\frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}\right]}{\partial\hat{\beta}}$, $Var(\hat{\beta}) = -[H(\hat{\beta})]^{-1}$ and $H(\hat{\beta}) = \frac{\partial^2 \sum_{i=1}^{N}\left[\frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}\right]}{\partial\hat{\beta}^2}$. Then, (1 - α) × 100% confidence interval is

$$\sum_{i=1}^{N}\left[\frac{\exp(x_{ij}\hat{\beta}_j)}{\sum_{k=1}^{J}\exp(x_{ik}\hat{\beta}_k)}\right] \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{Var(F_j)}, \quad (4)$$

where Φ^{-1}(.) is the inverse of cumulative distribution function of standard normal distribution. When α is 0.05, $\Phi^{-1}\left(1-\frac{\alpha}{2}\right) \approx 1.96$.

Case Study

Through the previous section, we have realized that the Delta method is essentially an approximation of the probability distribution using Taylor series expansion. The performance of the Delta method needs to be tested in a real dataset. In this section, a case study will be conducted to examine the performance of the Delta method in generating the confidence intervals of predictions from a multinomial logit model.

Data for the case study is extracted from 2000 Swiss Microcensus travel survey. A sample consisting of 3,408 commuting trips from the Aargau Canton is selected to estimate a multinomial logit model for travel mode choice. Four major commute modes are classified as auto, transit, bicycle and walk. 70% of trips (2,370 trips) are randomly selected for model estimation and the rest 30% (1,038 trips) are used for testing the confidence intervals of predictions.

Table 1 provides a descriptive analysis of variables in the sample for model estimation. Statistics of dummy variables indicating mode choices show the market shares: 60% of commuting trips use auto, 20% use transit, 11% use bicycle and 9% walk. The statistics well illustrate the multimodal transportation system in Switzerland. The lower part of Table 1 shows explanatory variables of travel times via alternative modes, which will be specified in the mode choice model. Table 2 shows similar statistics in the sample for test. Since these two samples are formed by randomly splitting one sample, statistics of choice variables and explanatory variables are very similar.

Table 3 provides the model estimation results of the MNL model based on 2,370 commuting trips. In the model, the alternative-specific constant for walk mode is fixed at 0. The alternative-specific constant is estimated at -0.1475 for auto mode, which is not quite significantly different from 0. The constants for transit and bicycle modes are -1.0893 and -0.3854. All the variables of travel times using alternative modes receive reasonably negative coefficients in the model.

The left part of Table 4 shows the confidence intervals and standard deviations (i.e. the square roots of the calculated variances) of predictions computed from the Delta method using Equation (4). The upper part of Table 4 shows the results within the sample for model estimation. As shown, 1,422 auto trips, 473 transit trips, 253 bicycle trips and 222 walk trips are observed in the sample. Since the model is estimated based on the same sample, the alternative-specific constants will be automatically adjusted to match the observed choice frequencies. As a result, the mean values calculated from Equation (2) perfectly match the observed counterparts.

The right part of Table 4 shows the simulation results based on 10,000 random draws of model coefficients. Model coefficients follow a multivariate normal distribution. We draw 10,000 sets of random numbers following multivariate normal distribution of model coefficients and compute 10,000 predictions. Then, mean values, standard deviations and 95% confidence intervals of predictions can be estimated and presented in the right part of Table 4. As shown, the theoretical results calculated from the Delta method are fairly close to those from simulation results. The slight difference is mainly caused by random errors in the simulation process. Due to random errors, the simulated mean values of choice frequencies are not exactly the same as the observed frequencies in the sample but the Delta method provides perfect mean values because they are calculated from analytical formulae.

The lower part of Table 4 compares theoretical and simulated results within the test sample. The test sample consists of 1,038 commuting trips, among which 637 trips use auto, 193 trips use transit, 114 trips use bicycle and 94 trips walk. The estimated model, as shown in Table 3, is applied to predict choice frequencies. The expected choice frequencies are calculated as 625, 203, 112 and 99 for four modes in sequence. As expected, they are not exactly the same as the observed counterparts because the model is not estimated based on the test sample. It demonstrates a good example for how to apply a discrete choice model to conduct prediction within a new sample. The variances of predicted choice frequencies are estimated by the Delta method and the standard deviations are 9.74, 8.05, 6.35 and 5.83 for four modes in sequence. Then, the 95% confidence intervals are calculated as the expected frequencies ± 1.96 × standard deviations, as listed in the table. It is not anticipated that the expected frequencies are exactly the same as the observed counterparts due to the nature of randomness but the observed counterparts are anticipated to fall into the computed confidence intervals. Otherwise, it will be considered as a small-probability event. In this case study, it can be seen that all the observed choice frequencies fall into the computed 95% confidence intervals, which are subject to the statistical principle. Figures 1 to 4 plot the histograms of simulated predicted choice frequencies for the test sample. As shown, all the histograms exhibit distributions close to normal distributions, which validates the theorem in the section of methodology.

The lower right part of Table 4 shows the similar simulated results. Again, due to random errors in simulation process, the mean values, standard deviations and simulated confidence intervals are just slightly different from the theoretical results given by the Delta method. The difference is negligible but the Delta method runs hundreds of times faster than the simulation method. Therefore, the Delta method should be highly recommended for computing confidence intervals of predictions from discrete choice models.

Conclusions and Discussions

In this paper, the Delta method is recommended to calculate variance and confidence interval of a prediction from discrete choice model. In a case study, the theoretical confidence intervals computed by the Delta method are compared with simulated counterparts and no obvious difference is found between them. This comparison has validated the Delta method through a real application. This paper also explicitly demonstrates how to apply a discrete choice model to conduct predictions in a scientific way. Discrete choice modeling method is developed based upon solid foundation of probability and statistical theory. It should be realized that a prediction from discrete choice model is not constant but random in nature.

At individual level, discrete choice model only provides a probability of an individual to choose an alternative, rather than the exact choice. The value of such a probability itself is also random in nature. Since the model coefficients are estimated based on a finite sample, estimated model coefficients are random and, in turn, the probability calculated based on those coefficients is also random. One may certainly use the Delta method to calculate the confidence interval of such a probability value if "N" takes value 1 in Formula 4. Nevertheless, we consider that prediction of choice frequencies at population level is of more practical significance. On one hand, it is extremely difficult to accurately predict an individual's travel choice behavior given the amount of information that is currently available. On the other hand, from the perspectives of travel demand modelers and transportation planners, a specific individual's travel choice will not affect the overall situation but an accumulation of many individuals' choices is the core of the travel demand forecast and transportation planning. Thus, it is of more concern about how to apply discrete choice model to accurately predict choice frequencies at population level. This paper introduces a technique for quantifying the accuracy level of predicted choice frequencies from a discrete choice model. It also provides a new perspective to evaluate the performance of discrete choice models based on the range of the confidence interval of predicted choice frequency. A narrow confidence interval of prediction indicates a high overall performance of a discrete choice model in terms of predictive power.

The future research effort will be made to apply the Delta method to compute confidence intervals for more complex models such as nested logit model, cross-nested logit model, multinomial probit, mixed logit model, etc. Conventionally, modelers only compare alternative models in terms of goodness-of-fit of data. Now we may think of comparing alternative models from a new perspective: the range of prediction's confidence interval. It may open some new interesting topics for future research.

References

Armstrong, P., R. Garrido, J. D. Ortuzar (2001). Confidence Intervals to Bound the Value of Time. Transportation Research Part E **37**, pp. 143 - 161.

Ben-Akiva, M. and S.R. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. the MIT Press, Cambridge, Massachusetts.

Bhat, C.R., A. Govindarajan, and V. Pulugurta (1998). Disaggregate Attraction-End Choice Modeling. Transportation Research Record **1645**, Transportation Research Board, National Research Council, Washington, D.C., 1998, pp. 60-68.

Cherchi, E. (2009). Modeling Individual Preferences, State of the Art, Recent Advances and Future Directions. Resource Paper Prepared for the Workshop on: Methodological Developments in Activity-Travel Behavior Analysis. The 12th International Conference on Travel Behavior Research, Jaipur, India, pp. 13-18.

Daly, A., S. Hess, G. Jong (2012). Calculating Errors for Measures Derived from Choice Modeling Estimates. Transportation Research Part B **46**, pp. 333–341

Gangrade, S., R.M. Pendyala and R.G. McCullough (2002). A Nested Logit Model of Commuters' Activity Schedules. Journal of Transportation and Statistics **5(2)**, pp. 19-36.

Greene, W. H. (2002). Econometric Analysis (5th Edition). Prentice Hall, Upper Saddle River, New Jersey.

Horowitz, J. (1979). Confidence Intervals for Choice Probabilities of the Multinomial Logit Model, Transportation Research Record **728**, pp. 23-29.

Mcfadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. in P. Zarembka, Frontiers in Econometrics, Academic Press, New York.

Pinjari, A.R., N. Eluru, C.R. Bhat, R.M. Pendyala and E. Spissu (2008). A Joint Model of Residential Location and Bicycle Ownership: Accounting for Self-Selection and Unobserved Heterogeneity. Transportation Research Record, Journal of the Transportation Research Board **2082**, pp. 17-26.

Train, K. (2009). Discrete Choice Methods with Simulation, second ed. Cambridge University Press, Cambridge, MA.

Ye, X. and R.M. Pendyala (2009). A Probit-based Joint Discrete-Continuous Model System: Analyzing Relationship between Timing and Duration of Maintenance Activities. Transportation and Traffic Theory **18**, pp. 403-423.

Zhang, Y. and Y. Xie (2008). Travel Mode Choice Modeling with Support Vector Machines. Transportation Research Record **2076**, pp. 141-150.

Table 1. Description of Data for Model Estimation
(Sample Size = 2,370)

Variables	Mean	Std. Dev.	Min.	Max.
Auto Dummy	0.60	0.49	0.00	1.00
Transit Dummy	0.20	0.40	0.00	1.00
Bicycle Dummy	0.11	0.31	0.00	1.00
Walk Dummy	0.09	0.29	0.00	1.00
Travel $Time_{Auto}$ (min)	14.74	12.28	0.17	81.00
Waiting $Time_{Transit}$ (min)	31.60	24.97	4.68	75.00
In-Vehicle $Time_{Transit}$ (min)	19.45	24.03	0.30	120.00
Travel $Time_{Bicycle}$ (min)	45.40	47.24	0.50	335.00
Travel $Time_{Walk}$ (min)	135.72	141.33	1.50	1005.00

Table 2. Description of Data for Test
(Sample Size = 1,038)

Variables	Mean	Std. Dev.	Min.	Max.
Auto	0.61	0.49	0.00	1.00
Transit	0.19	0.39	0.00	1.00
Bicycle	0.11	0.31	0.00	1.00
Walk	0.09	0.29	0.00	1.00
Travel $Time_{Auto}$ (min)	14.56	11.85	0.17	81.00
Waiting $Time_{Transit}$ (min)	31.65	25.13	4.68	75.00
In-Vehicle $Time_{Transit}$ (min)	20.33	25.53	0.30	120.00
Travel $Time_{Bicycle}$ (min)	45.43	47.40	0.50	309.00
Travel $Time_{Walk}$ (min)	135.57	141.60	1.50	928.00

Table 3. Model Estimation Results
(Sample Size = 2,370)

Variables	Est. Coeff.	Std. Dev.	T-test	p-value
$Constant_{Auto}$	-0.1475	0.1242	-1.1877	0.2350
$Travel\ Time_{Auto}$	-0.0442	0.0055	-8.0949	0.0000
$Constant_{Transit}$	-1.0893	0.1475	-7.3841	0.0000
$Waiting\ Time_{Transit}$	-0.0245	0.0031	-7.9146	0.0000
$In\text{-}Vehicle\ Time_{Transit}$	-0.0104	0.0044	-2.3578	0.0184
$Constant_{Bicycle}$	-0.3854	0.1395	-2.7631	0.0057
$Travel\ Time_{Bicycle}$	-0.0739	0.0053	-14.0404	0.0000
$Travel\ Time_{Walk}$	-0.0389	0.0028	-14.0097	0.0000
LL(β)	-2164.176			
LL(c)	-2580.372			
LL(0)	-3285.518			
$\rho^2(0)$	0.3413			
adj. $\rho^2(0)$	0.3389			
$\rho^2(c)$	0.1613			
adj. $\rho^2(c)$	0.1594			
Sample Size (N)	2370			

Table 4. Comparisons between Theoretical and Simulated Results

	Theoretical Results from Delta Method			Simulated Results		
Within the Sample for Model Estimation (N = 2,370)						
Commute Modes (Observed Numbers)	Mean	Std. Dev.	95% Confidence Interval	Mean	Std. Dev.	95% Confidence Interval
Auto (1422)	1422	22.29	[1378, 1466]	1420	22.11	[1376, 1463]
Transit (473)	473	18.48	[437, 509]	474	18.33	[439, 510]
Bicycle (253)	253	14.38	[225, 281]	254	14.24	[226, 282]
Walk (222)	222	13.14	[196, 248]	222	13.24	[197, 249]
Within the Sample for Test (N = 1,038)						
Auto (637)	625	9.74	[606, 644]	624	9.75	[605, 643]
Transit (193)	203	8.05	[187, 219]	203	8.03	[188, 219]
Bicycle (114)	112	6.35	[99, 124]	112	6.30	[100, 125]
Walk (94)	99	5.83	[87, 110]	99	5.83	[88, 111]

Figure 1. Histogram of Simulated Frequency of Auto Choices (N = 10,000)

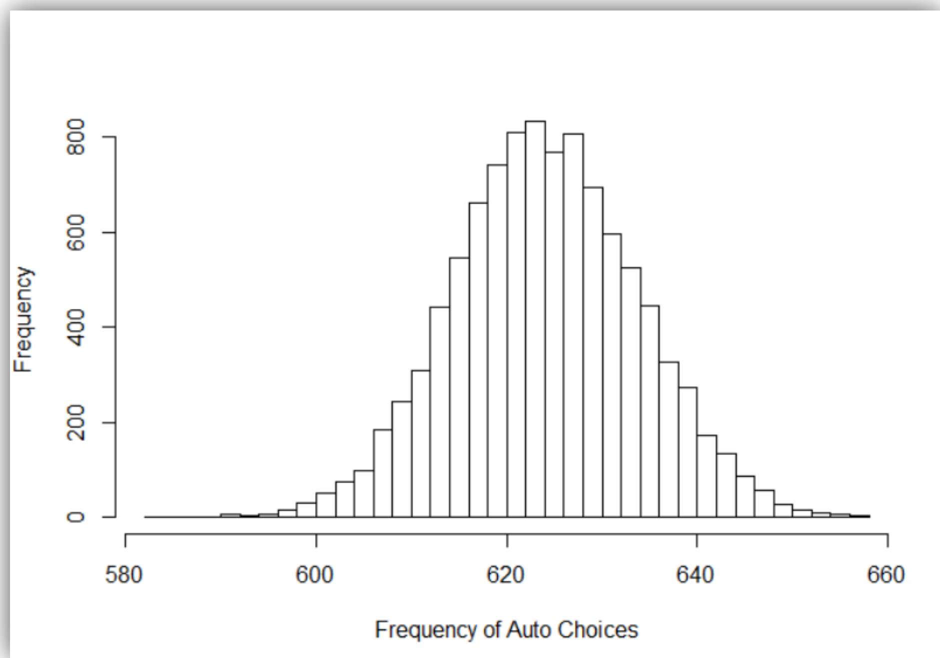

Figure 2. Histogram of Simulated Frequency of Transit Choices (N = 10,000)

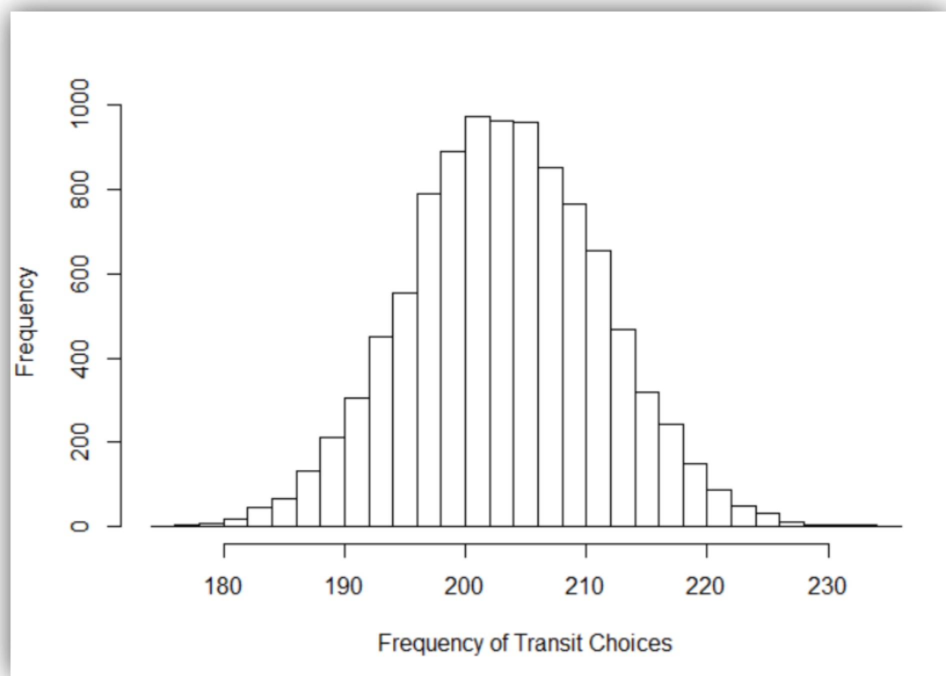

Figure 3. Histogram of Simulated Frequency of Bicycle Choices (N = 10,000)

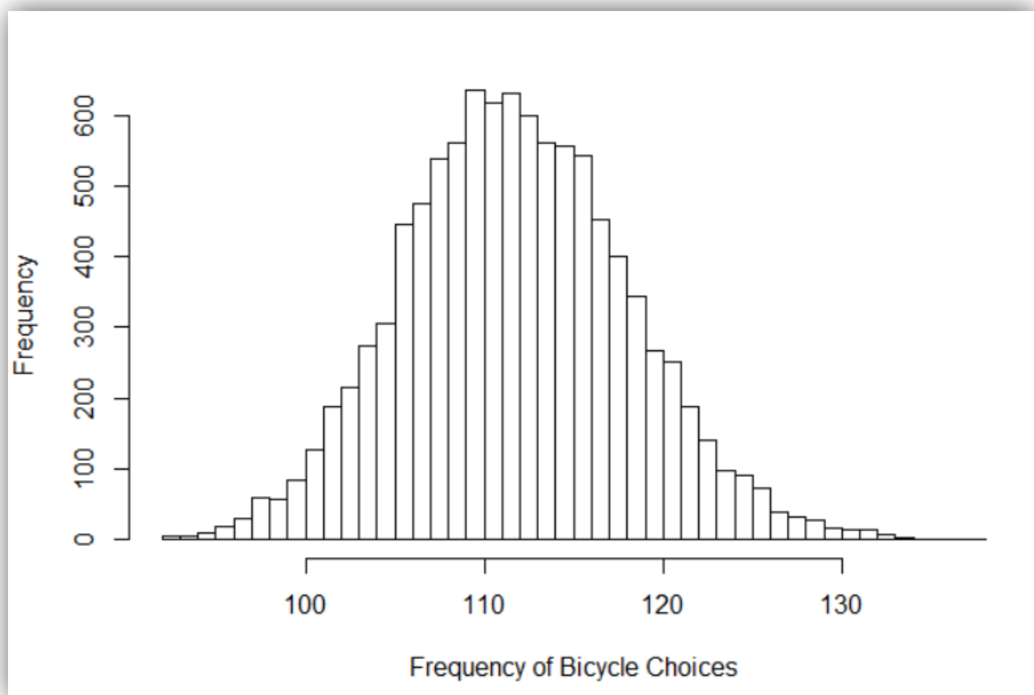

Figure 4. Histogram of Simulated Frequency of Walk Choices (N = 10,000)