IOP Conf. Series: Materials Science and Engineering **282** (2017) 012010 doi:10.1088/1757-899X/282/1/012010

A Study of Commuters' Decision-Making When Delaying Departure for Work-Home Trips

Fangjie Que and Wei Wang

School of Transportation, Southeast University, Nanjing 210096, China

E-mail:wangwei_transtar@163.com

Abstract. Studies on the travel behaviors and patterns of residents are important to the arrangement of urban layouts and urban traffic planning. However, research on the characteristics of the decision-making behavior regarding departure time is not fully expanded yet. In this paper, the research focuses on commuters' decision-making behavior regarding departure delay. According to the 2013 travel survey data of Suzhou City, a nested logit (NL) model was built to represent the probabilities of individual choices. Parameter calibration was conducted, so that the significant factors influencing the departure delay were obtained. Ultimately, the results of the NL model indicated that it performed better and with higher precision, compared to the traditional multinomial logit (MNL) model.

1. Introduction

Departure time prediction is an important component when studying commuters' travel decision-making behavior. Facing serious traffic congestion during daily peak periods, more people choose later departure time after work or school. Congestion charging, flexible schedules, and other traffic-demand management strategies were adopted in several cities, which provide commuters a wider range of choices in departure time and a greater utility of delaying departure. A discrete choice model was created in this paper, to help analyze the characteristics of resident travel behaviors and master the distribution regulation of departure time. The results can be used to predict travel behaviors and evaluate the effect of urban planning and management strategies.

2. Nested-Logit model construction method

The nested logit (NL) model, first derived by Moshe Benakiva [1], is a discrete choice model designed to capture correlation among alternatives. Unlike the traditional multinomial logit (MNL) model, it divides choice sets into nests, and the alternatives within each nest are correlated [2]. For example, in a two-level NL model, alternatives D are divided into nests M, based on the assumption that the alternatives in the same nest share the same variance of deviation [3]. Logically, individual n consider of choosing nest m, $m \in M$, before choosing the specific alternative d, $d \in D$, as the choice process.

Let the utility of the alternative dm chosen by commuter n be equation (1),

$$U_{dmn} = V_{dmn} + V_{mn} + V_{dn} + \varepsilon_{dmn} + \varepsilon_{mn} \quad (1)$$

where the deterministic components V_{dmn}, V_{mn}, and V_{dn} are defined as a linear combination of the measurable variables $\boldsymbol{X}_{dm}$, $\boldsymbol{X}_{m}$, and $\boldsymbol{X}_{d}$, with undetermined coefficients β_{dm}, β_{m}, and β_{d}, respectively. ε_{dmn} and ε_{mn} are random utilities.

It is assumed that ε_{dmn} follows a Gumbel distribution with scale parameter μ_d , while $\max_{d\in D} U_{dm}$ follows a Gumbel distribution with scale parameter μ_m . All people are assumed to choose the alternative providing the greatest utility.

Following the choice process mentioned above, the probability of the alternative *dm* chosen by commuter *n* is defined as a conditional probability shown as equation (2),

$$P_n(dm) = P_n(d \mid m)P_n(m) \tag{2}$$

The probability of the alternative *d* if nest *m* is chosen is expressed as equation (3),

$$P_n(d \mid m) = \frac{e^{(V_{dmn}+V_{dn})\mu_d}}{\sum_{d'\in D_{rn}} e^{(V_{d'mn}+V_{d'n})\mu_d}} \tag{3}$$

And the probability of nest *m* is **e**quation (4),

$$P_n(m) = \frac{e^{(V_{mn}+V^*_{mn})\mu_m}}{\sum_{m'\in M_n} e^{(V_{m'n}+V^*_{m'n})\mu_m}} \tag{4}$$

where V^*_{mn} is the logsum in equation (5),

$$V^*_{mn} = \frac{1}{\mu_d} ln \sum_{d\in D_{rn}} e^{(V_{dmn}+V_{dn})\mu_d} \tag{5}$$

These functions can be integrated into one equation and rewritten as equation (6),

$$P_n(dm) = P_n(d \mid m)P_n(m) = \frac{e^{(V_{dmn}+V_{dn})\mu_d}}{\sum_{d'\in D_{rn}} e^{(V_{d'mn}+V_{d'n})\mu_d}} \frac{e^{(V_{mn}+V^*_{mn})\mu_m}}{\sum_{m'\in M_n} e^{(V_{m'n}+V^*_{m'n})\mu_m}} \tag{6}$$

3. Departure delay choice modeling

3.1. Building selection tree

According to the commuter travel survey in the Central Business District of Suzhou, Jiangsu, the distribution of departure delay is shown in Figure 1.

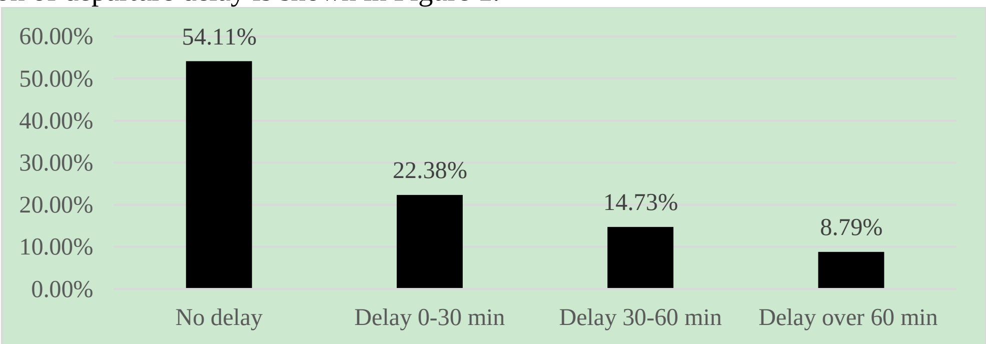

Figure 1. Histogram of the departure delay for work-home trips.

As can be seen from Figure 1, there are four available alternatives for departure delays for work-home trips. Nearly half of commuters delay departure for work-home trips. People who choose delay departures within 30 minutes comprise 22.38% of the population, while 8.79% of commuters delay departures greater than one hour. The alternatives can be partitioned into two nests *M*, *M*= {delay, no delay}. The former nest only includes one alternative, while the latter comprises three choices: delay

IOP Conf. Series: Materials Science and Engineering **282** (2017) 012010 doi:10.1088/1757-899X/282/1/012010

departure within 30 minutes, 30-60 minutes, and over one hour, numbered 12, 22, and 32, separately. The nested structure is shown in Table 1.

Table 1. The nested structure of the departure delay choice model.

Upper Level	Lower Level	Number	Percentage
No delay(1)	No delay (11)	665	54.11%
delay(2)	Delay 0-30 min(12)	275	22.38%
	Delay 30-60 min(22)	181	14.73%
	Delay over 60 min (32)	108	8.79%
Sum		1229	100%

3.2. Selecting variables
In this paper, nine factors were chosen as explanatory variables, after analysis of the survey data. The definitions and explanations of variables are listed in Table 2.

Table 2. Explanatory variables of the NL model.

Attribute	Variable	Variable Explanation
Individual	*gender*	1-male, 0-female
	age	Age
	income	Income
Household	*ftime*	If flexible work hours: 1-yes, 0-no
	car	Number of cars in household
	child	Number of children in household
Trip	*distance*	Distance from work to home (in miles)
	mode	Mode of transportation used work-to-home: 1-car SOV (single occupancy vehicle), 2-carpool, 3- vanpool, 4-bus,5-other.
Alternative-specific	*tatf*	Ratio of actual travel time to free-flow travel time

3.3. Parameter estimation
Data analysis and parameter estimation were completed using the Stata 12.0 application. With the base alternative of no delay, the final results are shown in Table 3.

Table 3. Final results of parameter estimation.

Lower level	Delay 0-30 min		Delay 30-60 min		Delay over 60 min	
constant	-7.77	(-3.53)	-11.95	(-3.90)	-13.33	(-5.53)
tatf	-13.05	(-7.02)	-13.05	(-7.02)	-13.05	(-7.02)
gender	—	—	—	—	-2.41	(-1.88)
age	0.06	(1.88)	-0.08	(-1.72)	—	—
ftime	—	—	2.29	(2.17)	—	—
distance	-0.35	(-2.59)	—	—	—	—
mode1	2.75	(2.40)	4.29	(2.60)	—	—
mode2	—	—	3.60	(1.78)	—	—
mode3	—	—	—	—	3.88	(2.19)
Upper level	Delay					

child	0.19	(2.78)
distance	0.28	(3.43)
mode4	2.17	(2.11)
Dissimilarity parameters:		
$\tau_{nodelay}$	1	—
τ_{delay}	0.67	4.45***
LR test for IIA (tau = 1):	*chi2(2) =9.79*	*Prob > chi2=0.007 5*

Near the end of the output, dissimilarity parameters exist, which measure the degree of correlation of random shocks within each of the two alternatives in the upper level. A likelihood-ratio test, with the null hypothesis that all parameters are one, can also be determined. Equivalently, the property known as the IIA [4], imposed by the multinomial logit model holds if and only if all dissimilarity parameters are equal to one [5]. The dissimilarity parameter in this model is 0.67, passing the significance test. Therefore, the null hypothesis is rejected, indicating that the alternatives in nest "delay" are not independent.

The equations of utility and *logsum* in the lower level can be determined by the results of the estimation in Table 3.

Utility values of alternatives in the lower level are given by:

V_{12}=-7.77-13.05*tatf*+0.06*age*-0.35*distance*-2.75*mode1,*

V_{22}=-11.95-13.05*tatf*-0.08*age*+2.29*ftime*+4.29*mode1*+3.60*mode2,*

V_{32}=-13.33-13.05*tatf*-2.41*gender*+3.88*mode3,*

and the *logsum* can be calculated based on equation (5):

V_2^*=0.67ln[exp(-7.77-13.05*tatf*+0.06*age*-0.35*distance*-2.75*mode1*)
+exp(-11.95-13.05*tatf*-0.08*age*+2.29*ftime*+4.29*mode1*+3.60*mode2*)
+exp(-13.33-13.05*tatf*-2.41*gender*+3.88*mode3*)].

The relative utility function of delay departure is expressed as:

V_2=0.19*child*+0.23*distance*+2.17*mode4*.

Ultimately, the probabilities of departure delay alternatives were computed by the expressions in Table 4.

Table 4. Probability expressions of each departure delay alternative for work-home trips.

Upper Level	Lower Level	Probability
No delay (1)	No delay (11)	$\frac{1}{1+\exp(-V_2-V_2^*)}$
Delay (2)	Delay 0-30 min (12)	$\frac{1}{1+\exp(V_2+V_2^*)}\cdot\frac{\exp(V_{12})}{\exp(V_{12})+\exp(V_{22})+\exp(V_{32})}$
	Delay 30-60 min (22)	$\frac{1}{1+\exp(V_2+V_2^*)}\cdot\frac{\exp(V_{22})}{\exp(V_{12})+\exp(V_{22})+\exp(V_{32})}$
	Delay over 60 min (32)	$\frac{1}{1+\exp(V_2+V_2^*)}\cdot\frac{\exp(V_{32})}{\exp(V_{12})+\exp(V_{22})+\exp(V_{32})}$
	Sum	1

4. Discussion

4.1. Bottom model parameters

Table 3 shows that there are eight factors significantly influencing the decision-making in the lower level.

Considering the variables with individual attributes, gender only affects the choice of delaying departure for more than one hour and, on equal conditions, women are more likely to delay departure for more than 60 minutes than men. The estimates of age were 0.062, -0.079, -, indicating that the elderly prefer not to go home too late. Flexible schedules provide more choices for departure time, and the utility of choosing a 30-60 minutes' delay significantly increases.

The family characteristics variables, including number of children and cars in household, have little effect on the length of delay.

Commuter distance influences when people choose a departure delay of less than 30 minutes. Regarding the mode of transportation used from work-to-home, car SOV, carpool, and vanpool affect the alternative No. 12, 22, and 32, respectively, if the alternative of delay were chosen. Consequently, reasonable resource allocation of both time and space must be provided when developing and promoting ride sharing.

Of all the factors in the lower model, the most significant factor affecting the departure time is the ratio of the actual travel time to free-flow travel time, which is a determinant for commuters when making decisions on the departure time for work-home trips.

4.2. Upper model parameters

The number of children in the household, distance from work to home, and use of buses are the main factors in the upper model. Commuters with more children are more likely to depart on time. People with shorter commute distances tend to delay departure. Therefore, developing new towns and constructing polycentric urban structures allow people to choose a wider variety of work places and make diversified schedules. For commuters who choose buses as their transport, the utility value of delay will be greatly reduced. Hence, a reasonable arrangement of bus schedules is important to commuters' decision-making regarding departure time.

4.3. Prediction accuracy

To evaluate the model, the accuracy of model prediction can be described by two indexes: sensitivity and specificity. The receiver operating characteristic curve (ROC) is a comprehensive evaluation of the sensitivity and specificity of the established model, utilizing geometric methods to reveal the relationship between the two indexes [6]. Generally, the value of the area under the curve (AUC) is between 0.5 and 1. The closer the area is to 1, the higher the prediction accuracy of the model. In this paper, the ROC curve of the NL model is shown in Figure 2.

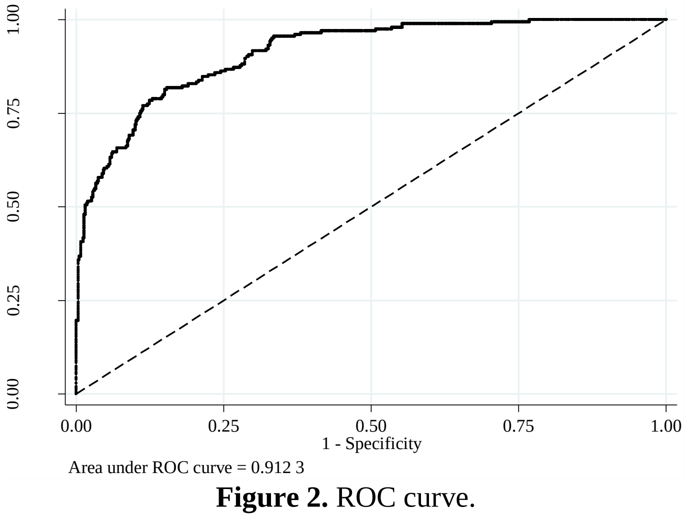

Figure 2. ROC curve.

The area under the ROC curve (AUC) is 0.9123, which indicates the model prediction has high accuracy.

5. Conclusion

The principle and formula derivation of the NL model is introduced in this paper, and is applied to the study of decision-making regarding departure delays. After analysis of the characteristics of commuters' delays on work-home trips, the structure of the delay alternative tree is provided. According to survey data in Suzhou City, the parameter calibration is conducted via Stata. The results show that the model has high accuracy, and describes the characteristics of commuters' decision-making regarding delaying departure on work-home trips better than the traditional MNL model, overcoming the defects of the independence of irrelevant alternatives, which has a reference and adaptive significance.

To acknowledge the travel time characteristics more comprehensively and scientifically, and understand more practical approaches to solving decision-making behavior problems, further research could be performed in the following respects. First, the quantity of explanatory variables should be enriched to improve the model, adding cost and level of service of roads as independent variables. Next, the model can be applied to the evaluation of traffic management measurements, such as the before-after study of congestion pricing. Also, the model could be used for urban traffic demand forecasting and transportation planning evaluation, which has distinctly practical value.

References

[1] Moshe B 1974 *J*. Structure of passenger travel demand models. *Transportation Research Board Record*. **526** pp 26-42
[2] Daniel M 1974 Conditional logit analysis of qualitative chioce behavior *Frontiers in Econometrics* ed Zarembka P (New York: Academic Press) pp 105–142
[3] Moshe B and Steven R L 1985 *Discrete choice analysis: theory and application to travel demand* (London: MIT Press)
[4] Hensher D and Greene W 2002 *J*. Specification and estimation of the nested logit model: alternative normalizations. *Transportation Research Part B: Methodological*. **36(1)** 1-17
[5] Tuansheng C 2007 Travel behavior characteristics analysis for commuter. Doctoral Thesis. (Beijing: Beijing Jiaotong University)
[6] Mario A C 1999. Receiver operating characteristic (ROC) analysis. *Stata Technical Bulletin* **52** pp19-33