Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 25 (2017) 2649–2659

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016

Effects of cost adjustment on travel mode choice: analysis and comparison of different logit models

Tongzhou Bai[a,b,]*, Xuemei Li[a], Zhiyuan Sun[c]

[a]*School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China*
[b]*Beijing Transportation Research Center, Beijing 100073, China*
[c] *Institute of Transportation Engineering, Tsinghua University, Beijing 100084, China*

Abstract

Residence mode choice behavior is affected by various factors, and research on the effects of travel cost would give valuable guidance in making cost-related travel demand management policy. Logit models are widely used in mode choice analysis and this paper estimates three kinds of logit models, namely MNL, NL and ML, to choose the most suitable one in analyzing cost adjustment affection. Estimation results show that market shares of car and PT interact obviously given cost change. 10% decrease of PT cost seems to have a better promoting effect of PT share increase than 10% increase of car cost.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: travel behavior analysis; travel mode choice; Logit model; travel cost; model comparison

1. Introduction

Along with the rapid social and economic development and the adjustment of policy guidance, urban travel modal split changes tremendously. It has been widely accepted that research on travel mode split and mode choice preference is of great significance in urban planning and management. Residence mode choice behavior is affected by various aspects, such as parking fee, activity demand and purpose, urban land use, etc., among which travel cost is an influential factor. Research on the potential effects of travel cost would give guidance in making cost-related travel demand management policy.

* Corresponding author. Tel.: +86-10-57079793; fax: +86-10-57079800.
E-mail address: tongzhou_bai@163.com

2352-1465
Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.
10.1016/j.trpro.2017.05.150

Traditionally discrete choice models are implemented in analyzing travel behavior, among which logit model is frequently used in a great deal of literatures. Logit model is the name of a model group comprised of multinomial logit model, nested logit model, mixed logit model, etc. Zhang, Dai, Zhang (2012) use multinomial logit model to build a residence travel mode choice model considering walking surroundings, travel time, travel cost and based on utility equivalent principle to evaluate the walking affect area of rail transit station. Zhang, Guan, Qin et al. (2014) design SP survey to capture the factors effecting car and public transit travel in urban area and use MNL to investigate the factors that can enhance public transit attraction. Liu, Zhou (2006) give an overall introduction of how nested logit model is applied in travel mode choice evaluation. Zhang, Yao, Yang (2014) constructed nested logit model combing travel mode choice and route choice to investigate traffic flow assignment method under the condition of multi-modal. Luo, Sun, Wu (2010) discuss the shortcomings of traditional disintegrated logit model and apply mixed logit model to predict traffic mode split under diverse congestion pricing schemes. Wu, Lu, Wang et al. (2014) induce loyalty variable into the utility function and build a mixed logit model to investigate the mode choice behavior of intercity travels.

Most literatures focus on evaluation and application of one model category, though few compare three kinds of models. As to MNL, it has strict requirement of the alternatives to obey the independence of irrelevant alternatives (IIA) feature. In reality, affected by the unknown part of the utility function and relevance among alternatives, MNL often turns out to make fuzzy predictions. Single logit model may have shortcomings under different circumstances, so this paper attempt to estimate three kinds of logit models, namely MNL, NL and ML, to choose the most suitable one in analyzing cost adjustment affection.

In Section 2, three kinds of logit models are constructed and estimation method is interpreted. Section 3 explains the estimation algorithm of the models. Section 4 display the estimation result and discussion of what the estimates indicate. The most appropriate model for policy analysis is chosen based on the estimation results. In Section 5, the chosen model is applied to analyze the effects of cost adjustment on residence travel mode choice. Section 6 concludes that polices decreasing the cost of public transport tend to have better result than increasing car travel cost by the same proportion.

2. Methodologies

2.1. Multinomial logit model

Multinomial Logit Model (MNL) is widely used in analyzing travel related choice problem, such as travel mode choice, route choice, etc. MNL sign utility to each alternative *i* for decision maker *n*, noted as U_{ni} , and the utility function is composed of two parts, namely the deterministic part and the random part.

$$U_{ni}=V_{ni}+\varepsilon_{ni} \tag{1}$$

in which, V_{ni} is the utility that can be precisely described and measured, such as travel time, waiting time, car ownership, etc. ε_{ni} is the unpredicted part of the utility, and it is independent from the unknown part of other alternative and follows the Gumbel distribution.

The process of constructing MNL is basically a trade-off of variables in the utility function of different modes. Two important principles are considered in choosing variables, as shown below.

- Practically meaningful. This means to include variables that are important in real life and are easily interpretable in a practical way.
- Statistically meaningful. In the process of changing variables, the previous estimation is considered as a guide. The sign and relative magnitudes, as well as the t-value which representing statistical significance are vital in choice of certain variables.

To make it easier for estimation, alternative trip modes are restricted to four categories, and the choice set is M={walk, bike, car, public transport}. A multinomial logit model is applicable on the assumption of independence of irrelevant alternatives (IIA).

The data analysed in this paper is taken from the travel survey conducted in Stockholm region, containing only weekday trips. Each trip data contains 33 variables and up to 21500 trips are recorded.

Two specifications of MNL are constructed; specific 1 has more variables including auxiliary explanative variables and specification 2 has fewer variables but mainly basic ones like travel time, distance, etc. The utility functions (the deterministic part) are shown below.

Specification 1:

$$\begin{cases} V_{walk} = ASC_{walk} + \beta_{walk\_distance} \cdot Distance + \beta_{walk\_dist<2km} \cdot I_{dist<2km} \\ V_{bike} = ASC_{bike} + \beta_{bike\_distance} \cdot Distance + \beta_{bike\_dist<2km} \cdot I_{dist<2km} + \beta_{bike\_male} \cdot I_{male} \\ +\beta_{bike\_cons\_cap} \cdot Cons\_cap + \beta_{bike\_househ\_w\_kids} \cdot I_{househ\_w\_kids} + \beta_{bike\_flext} \cdot I_{flext} \\ V_{car} = ASC_{car} + \beta_{car\_time} \cdot Cartime + \beta_{car\_cost} \cdot Carcost + \beta_{car\_compcar} \cdot I_{compcar} \\ +\beta_{car\_male} \cdot I_{male} + \beta_{car\_dgt20km} \cdot I_{dist>20km} + \beta_{car\_own\_drv\_lic} \cdot I_{own\_drv\_lic} \\ +\beta_{car\_cheappark} \cdot I_{cheappark} + \beta_{car\_cons\_cap} \cdot Cons\_cap + \beta_{car\_househ\_w\_kids} \cdot I_{househ\_w\_kids} \\ +\beta_{car\_flext} \cdot I_{flext} \\ V_{PT} = \beta_{PT\_time} \cdot PTtime + \beta_{PT\_cost} \cdot PTcost + \beta_{PT\_noboard} \cdot PTnoboard \\ +\beta_{PT\_male} \cdot I_{male} + \beta_{PT\_cons\_cap} \cdot Cons\_cap + \beta_{PT\_househ\_w\_kids} \cdot I_{househ\_w\_kids} \\ +\beta_{PT\_flext} \cdot I_{flext} + \beta_{PT\_fwt} \cdot PTfwt \end{cases} \tag{2}$$

Specification 2:

$$\begin{cases} V_{walk} = ASC_{walk} + \beta_{walk\_distance} \cdot Distance + \beta_{walk\_dist<2km} \cdot I_{dist<2km} \\ V_{bike} = ASC_{bike} + \beta_{bike\_distance} \cdot Distance + \beta_{bike\_dist<2km} \cdot I_{dist<2km} \\ +\beta_{bike\_cons\_cap} \cdot Cons\_cap \\ V_{car} = ASC_{car} + \beta_{car\_time} \cdot Cartime + \beta_{car\_cost} \cdot Carcost \\ +\beta_{car\_parkposs} \cdot I_{parkposs} + \beta_{car\_male} \cdot I_{male} + \beta_{car\_dgt20km} \cdot I_{dist>20km} \\ +\beta_{car\_own\_drv\_lic} \cdot I_{own\_drv\_lic} + \beta_{car\_md} \cdot I_{car\_md} + \beta_{car\_income} \cdot Income \\ +\beta_{car\_cons\_cap} \cdot Cons\_cap \\ V_{PT} = \beta_{PT\_time} \cdot PTtime + \beta_{PT\_cost} \cdot PTcost + \beta_{PT\_innercity} \cdot I_{innercity} + \beta_{PT\_card} \cdot I_{PTcard} \\ +\beta_{PT\_cons\_cap} \cdot Cons\_cap \end{cases} \tag{3}$$

The interpretations of used variables are shown in Table 1. More detailed discussion of variables will be presented later along with the estimation outcome.

Table 1. Interpretation of included variables in the utility function

Variable Name	Parameter	Description
ASC	Alternative specific constant	Take public transport as reference
$Distance$	Distance	Trip distance(in mil=10 km)
$I_{dist<2km}$	Distance	1 if distance is less than 2km
$I_{dist>20km}$	Distance	1 if distance is more than 20 km
$Income$	Income	Monthly income in SEK. 1:0-7500, 2:7501-10000, 3:10001-15000, 4:15001-25000, 5:25001-40000, 6:40001-55000, 7:55001-70000, 8:70001+
$Cons\_cap$	Consumption capacity	Consumption capacity according to the social norm. 1:low, 2:low to average, 3:average, 4:average to high, 5:high
$Cartime$	Car time	Car driving time(in hour)
$Carcost$	Car cost	Car cost based on per distance cost, e.g. fuel(16 SEK/mil)
$PTtime$	Public transport time	Total travel time by public transport
$PTcost$	Public transport cost	Cost of public transport trip
$PTnoboard$	Public transport transfer	Public transport: number of boardings
$PTfwt$	Public transport time	Public transport: first waiting time
I_{male}	Male	1 if male
$I_{innercity}$	Inner city household location	1 if residence located in the inner city
$I_{househ\_w\_kids}$	Children	1 if household with kids 0-12 years
$I_{own\_drv\_lic}$	Driving license	1 if respondent has driving license
I_{flext}	Flextime working	1 if flexible working hours
$I_{parkposs}$	Work place parking	1 if parking possibilities at workplace
$I_{cheappark}$	Work place parking	1 if free or cheap parking at workplace
$I_{compcar}$	Company car	1 if access to company car
I_{car_md}	Car accessibility daily	1 if car accessible on measurement day
I_{PTcard}	Public transit card	1 if possessing travel card for public transport on measurement day

In MNL, the error term of utility function is assumed to be independently, identically distributed with Gumbel distribution. And the probability of choosing specific mode is given by,

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \tag{4}$$

In which, n represent the decision maker and j represent a specific mode in the mode choice set.

2.2. Nested logit model

If alternatives in the choice set can be divided into subsets (noted as nests) by certain criteria, the nested logit model can be constructed on the basis of the following properties:

- For any two alternatives that are in the same nest, the ratio of probabilities is independent of the attributes or existence of all the other alternatives. That is, IIA holds within each nest.
- For any two alternatives in different nests, the ratio of probabilities can depend on the attributes of other alternatives in the two nests. IIA does not hold in general for alternatives in different nests.

Walk and bike are relatively slow travel modes, while car and bus are faster modes. They are both correlated with each other, and can be described by nested logit model. To measure the degree of independence of the unobserved utility among the alternatives in nest *k*, the parameter λ_k is introduced. Technically, there are three ways to construct the model; the first is to set up one nest of {car, public transport} or {walk, bike}; the second is to build nests for {car, public transport} and {walk, bike}, with the same λ ; the third is to build two nests with different λ_k for each nest.

By comparison, the paper chooses to use the third way described above to construct the nested model, which interprets most precisely the characters of nests. The structure is shown below.

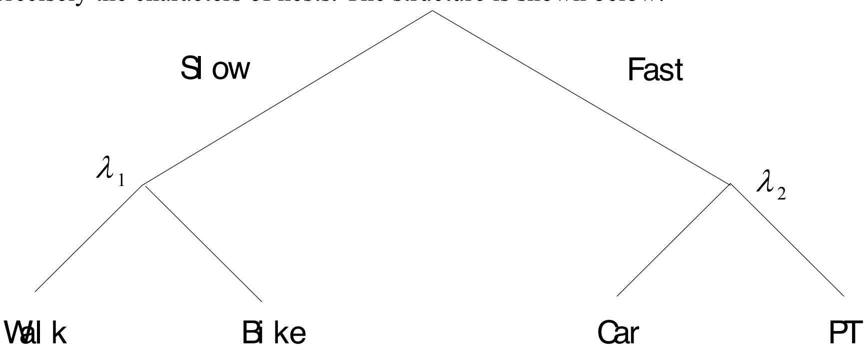

Figure 1 The structure of nested logit model

In terms of model specification, we choose to use the same utility function as specification 1 in MNL. The probability for alternative $i \in B_k$:

$$P_{ni} = \frac{e^{V_{ni}/\lambda_k} (\sum_{j \in B_k} e^{V_{nj}/\lambda_k})^{\lambda_k - 1}}{\sum_{l=1}^{K} (\sum_{j \in B_l} e^{V_{nj}/\lambda_l})^{\lambda_l}} \tag{5}$$

in which, B_k represents the nest to which a specific alternative belong.

2.3. Mixed logit model

Mixed logit model obviates the three limitations of standard logit by allowing for random taste variation, unrestricted substitution patterns and correlation in unobserved factors over time. One widely used application of mixed logit is based on random coefficients, which means that the coefficients vary over decision makers in population with density $f(\beta)$. When implementing the mixed logit model, the parameter distribute with a density instead of being fixed.

The purpose of using mixed logit model is basically to capture variety of parameter within population. Intuitively, travel cost in practice varies with different groups, such as income group. People with different income may have fiercely different sensitivity to the travel cost. So this paper chooses to mix the coefficient of public transport cost.

Generally, for a cost variable, we will always expect it to have a negative sign in coefficient, since people will tend to lose utility if the alternative increases in its cost. In terms of distribution, lognormal distribution is used in this paper, since it is useful when the coefficient is known to have the same sign for every decision maker.

Apart from mixing one variable in the utility, we use the same utility function as model specification 1 in MNL. The choice probability of mixed logit model is

$$P_{ni} = \int L_{ni}(\beta_{PT\_\cos t}) f(\beta_{PT\_\cos t} \mid \theta) d\beta_{PT\_\cos t} \tag{6}$$

where

$$L_{ni}(\beta) = \frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{nj}}} \tag{7}$$

3. Estimation algorithms

3.1. Maximum likelihood algorithm

Maximum likelihood estimation is widely used in estimating the parameters of a statistical model. The probability of a decision maker *n* choosing an alternative that s/he actually chooses is $\prod_i (P_{ni})^{y_{ni}}$, in which $y_{ni} = 1$ if *n* choose *i* and zero otherwise.

Assume that every decision maker choose alternative independently, the probability of all decision maker choosing their actual choices is

$$L(\beta) = \prod_{n=1}^{N} \prod_i (P_{ni})^{y_{in}} \tag{8}$$

in which β is a vector of the parameters to be estimated. The log-likelihood function is

$$LL(\beta) = \sum_{n=1}^{N} \sum_i y_{ni} \ln P_{ni} \tag{9}$$

the estimate result is then the β that maximizes this function, that is the derivative with respect to β is zero.

$$\frac{dLL(\beta)}{d\beta} = 0 \tag{10}$$

3.2. Mixed logit model estimation

To estimate the mixed logit model, traditional maximum likelihood algorithm is slightly changed into a simulated way. In mixed logit model, the probabilities are approximated through simulation for any given value of θ:

- Draw a value of β from $f(\beta_{PT\_\cos t} \mid \theta)$, and label it β^r with the superscript *r=1* referring to the first draw.

- Calculate the logit formula $L_{ni}(\beta^r)$ with this draw.
- Repeat steps 1and 2 many times, and average the results.

This average is the simulated probability:

$$\check{P}_{ni} = \frac{1}{R}\sum_{r=1}^{R} L_{ni}(\beta^r) \tag{11}$$

The simulated probability is inserted into the log-likelihood function to give simulated log likelihood:

$$SLL = \sum_{n=1}^{N}\sum_{j=1}^{J} d_{nj} \ln \check{P}_{nj} \tag{12}$$

4. Results and discussion

4.1. Results of MNL

On the basis of the travel data collected in Stockholm region, the estimation results of parameters are shown in the table below, along with the log-likelihood of the model and estimates of VoT (value of time).

Table 2. Estimation of two specification of MNL

Parameter	Model specification 1		Model specification 2	
	Estimate	t-value	Estimate	t-value
ASC_{walk}	1.45	4.27	2.01	6.14
ASC_{bike}	**0.411**	**1.62**	0.71	3.02
ASC_{car}	-0.586	-2.3	-0.786	-2.81
$\beta_{walk\_distance}$	-7.78	-14.4	-9.18	-14.9
$\beta_{bike\_distance}$	-2.2	-15.5	-2.19	-15.6
$\beta_{walk\_dist<2km}$	1.08	6.04	0.738	4.05
$\beta_{bike\_dist<2km}$	0.748	4.57	0.813	5.02
β_{car_time}	-1.03	-17.7	-0.482	-6.53
β_{PT_time}	-0.968	-3.27	-1.38	-5.66
β_{car_cost}	-0.0215	-3.84	-0.0309	-5.65
β_{PT_cost}	-0.0471	-5.87	-0.0388	-4.97
β_{bike_male}	0.337	2.49		
β_{car_male}	0.39	2.97	0.253	3.68
β_{PT_male}	**-0.047**	**-0.352**		
$\beta_{bike\_cons\_cap}$	-0.149	-2.7	-0.198	-3.9
$\beta_{car\_cons\_cap}$	-0.209	-3.87	-0.145	-2.72
$\beta_{PT\_cons\_cap}$	-0.114	-2.08	-0.159	-3.13
$\beta_{bike\_househ\_w\_kids}$	0.517	3.27		
$\beta_{car\_househ\_w\_kids}$	**0.117**	**0.745**		
$\beta_{PT\_househ\_w\_kids}$	0.469	2.94		
β_{bike_flext}	-0.37	-2.71		

β_{car_flext}	-0.495	-3.75		
β_{PT_flext}	-0.304	-2.27		
$\beta_{car\_compcar}$	0.856	6.83		
$\beta_{car\_dgt20km}$	0.324	2.12	0.426	2.81
$\beta_{car\_own\_drv\_lic}$	1.3	9.52	0.575	4
$\beta_{car\_cheappark}$	1.09	15.6		
$\beta_{car\_parkposs}$			0.681	9.61
β_{car_md}			1.77	15.9
β_{car_income}			-0.125	-4.21
β_{PT_fwt}	-1.3	-1.96		
$\beta_{PT\_noboard}$	**-0.0629**	**-0.986**		
$\beta_{PT\_innercity}$	0.53	6.1	0.51	5.92
β_{PT_card}	3.1	34.1	3.18	35.2
Number of estimated parameters	31		21	
Log-likelihood	-4702.643		-4681.061	
Log-likelihood for zero beta:	-11348.206		-11348.206	
McFadden rho:	0.586		0.588	
Log-likelihood for constants only:	-8964.125		-8964.125	
VoT_Car(SEK/h)	48.1		15.6	
VoT_PT(SEK/h)	20.5		35.6	

As stated in the former section, the process of choosing inclusive variables is a trade-off regarding both practical meaning and statistical efficiency. Apart from the basic variables such as distance, cost and time, this paper includes some more decision maker related variables like gender, consumption capacity and so on. In order for easier comparison, the walk mode is taken as reference. Intuitively, consumption capacity is more important than income when people make choices, since it represents the actual disposable money and besides, during the estimation process, the income data cannot give satisfactory estimation.

For model specification 1, number of children in the household is included, for this in reality is a very important consideration when parents make choices on travel modes. The flexibility of work time implies the sensitivity to travel time, which also may dominate people's choice of mode. The number of boarding, which implies the number of transfers during a trip, has a bad statistical efficiency, but it has an expected sign, which keeps it in the model. Gender coefficient for public transport and number of children in the household coefficient for car has bad statistical efficiency, but in order for comparison between different modes, both variables are kept in the model. ASC for bike is insignificant, but since only the relative magnitude is considered, even if it is not significantly different from 0, its value is between the values of walk and car. This does not affect the comparison result.

In model 2, some insignificant variables are removed and others such as parking possibility indicator and company car indicator are included.

In terms of comparison between the two models, all the coefficients have the expected signs. Model 1 has some insignificant coefficients, like ASC_{bike}, β_{PT_male}, $\beta_{car\_househ\_w\_kids}$, $\beta_{PT\_noboard}$, while model 2 has none. The McFadden rhos of the two models are almost the same, but the numbers of variables are different. By Chi-square test, we can judge that the two models are significantly different from each other at 95% level. When it comes to the value of time, the values in model 1 is more reasonable than model 2, since in practice car users tend to have higher value of time than public transport users. In a word, though model 2 seems to be better than model 1 in terms of statistical testing, we choose to propose model 1, since it gives more reasonable value of time and besides its shortcoming in statistics is acceptable.

4.2. Results of NL and Mixed logit

The estimation of parameters for nested logit model and mixed logit model are shown in the table below.

Table 3. Estimation of nested logit model and mixed logit model

Parameter	Nested logit model		Mixed logit model	
	Estimate	t-value	Estimate	t-value
ASC_{walk}	1.41	4.5	1.46	4.36
ASC_{bike}	**0.476**	**1.85**	**0.406**	**1.61**
ASC_{car}	-0.526	-2.19	-0.628	-2.47
$\beta_{walk\_distance}$	-6.91	-6.16	-7.6	-14.5
$\beta_{bike\_distance}$	-2.25	-13.2	-2.19	-15.5
$\beta_{walk\_dist<2km}$	1.07	5.77	1.01	5.73
$\beta_{bike\_dist<2km}$	1.03	5.17	0.7	4.3
β_{car_time}	-0.919	-13.9	-1.02	-17.5
β_{PT_time}	-1.03	-3.74	-0.979	-3.31
β_{car_cost}	-0.0205	-3.89	-0.0203	-3.63
β_{PT_cost}	-0.0392	-4.88		
β_{bike_male}	0.233	1.95	0.339	2.52
β_{car_male}	0.311	2.54	0.389	2.98
β_{PT_male}	**-0.0884**	**-0.712**	**-0.0456**	**-0.344**
$\beta_{bike\_cons\_cap}$	-0.118	-2.29	-0.136	-2.48
$\beta_{car\_cons\_cap}$	-0.172	-3.35	-0.189	-3.53
$\beta_{PT\_cons\_cap}$	**-0.0981**	**-1.9**	**-0.0965**	**-1.79**
$\beta_{bike\_househ\_w\_kids}$	0.353	2.28	0.481	3.07
$\beta_{car\_househ\_w\_kids}$	**0.0444**	**0.292**	**0.117**	**0.754**
$\beta_{PT\_househ\_w\_kids}$	0.32	2.07	0.434	2.75
β_{bike_flext}	-0.319	-2.59	-2.62	2.53
β_{car_flext}	-0.447	-3.6	**-3.72**	**0.4**
β_{PT_flext}	-0.268	-2.14	**-2.3**	**0.832**
$\beta_{car\_compcar}$	0.779	6.53	0.851	6.81
$\beta_{car\_dgt20km}$	0.295	2.12	**0.264**	**1.72**
$\beta_{car\_own\_drv\_lic}$	1.18	8.58	1.29	9.43
$\beta_{car\_cheappark}$	0.988	13.2	1.09	15.7
β_{PT_fwt}	**-0.874**	**-1.45**	**-1.14**	**-1.73**
$\beta_{PT\_noboard}$	**-0.0368**	**-0.631**	**-0.0379**	**-0.594**
$\beta_{PT\_innercity}$	0.523	6.5	0.524	6.04
β_{PT_card}	2.86	21.6	3.07	34
λ_1	0.732	3.66		
λ_2	0.889	19.4		
mu			-3.01	-18.8
sigma			0.0631	0.203
Number of estimated parameters	33		32	
Log-likelihood	-4701.499		-4703.405	
Log-likelihood for zero beta:	-11348.206		-11348.206	
McFadden rho:	0.586		0.586	
Log-likelihood for constants only:	-8964.125		-8964.125	
VoT_Car(SEK/h)	44.9		49.9	
VoT_PT(SEK/h)	26.3		19.8	

Since the same utility function as specification 1in MNL is used when estimating nested logit and mixed logit, the descriptive variables are the same in the two kinds of models. The only difference is that in nested logit two new parameters indicating the degree of independence of the nests are included. And in mixed logit model two new parameters showing the character of density are included, while the coefficient of $\beta_{PT\_\cos t}$ for the convenience of estimation is removed. The value of $\beta_{PT\_\cos t}$ could be replace by its mean value when calculate the value of time, which in this case is calculated as $mean = -\exp(mu + sigma \wedge 2/2) = -0.0494$.

Seen from the nest parameter, alternatives in the second nest {car, PT} have greater independence and less correlation than alternatives in the first nest {walk, bike}. But none of the nest parameters is equal to 1, which means the endeavour to build nested logit model is meaningful to some extent.

The newly included parameter in mixed logit model shows the density of the mixed parameter. The variance of coefficient can be calculated as: $\text{var}\,iance = \exp(2mu + sigma \wedge 2)(\exp(sigma \wedge 2) - 1) = 9.7E - 6$, which shows that the variance is very small and we can simply treat the coefficient of public transport cost as fixed among population.

The values of coefficients estimated by these two models are almost the same, except that mixed logit model have three more insignificant coefficients. The two models have the same McFadden rho and one different parameter number. So in general the two models are of the same quality. Since the number of insignificant estimation of nested logit model is fewer than that of mixed logit model, from a statistical view of points, we will choose the nested logit model as a better one between the two.

To compare the nested logit model with the chosen specification of MNL, they have the same McFadden rho. And nested model have two more parameters as well as more insignificant coefficients. In this sense, this paper chooses the model specification 1 of MNL as the best estimated model and will do the policy analysis based on it in the latter section. The decision can be supported by the comparison of model prediction tables.

Table 4. Model prediction table for the best model

	Walk	Bike	Car	PT
Walk	0.354	0.155	0.0341	0.0211
Bike	0.2087	0.233	0.0779	0.0477
Car	0.1691	0.3721	0.7286	0.1422
PT	0.2682	0.24	0.1594	0.7889

Column is the actual chosen mode, while row is the predicted choice and the numbers show the precision of prediction. The predictions of car and PT are more accurate than walk and bike. This same applies to the other models and this one is slightly more precise.

5. Cost adjustment analyses

Big cities around the world like Stockholm and Beijing have made plans to encourage the development of public transport while limit the usage of private cars. In Stockholm region, congestion charge was launched since 2006. Beijing has allocated plenty of budgets into construction of rail transit system. Among all the measures promoting urban transport development, economic measures have long been known as effective and policies like congestion charging, parking fee adjustment, etc. have been applied world widely.

When designing economic policy schemes, travel cost adjustment is considered. Thus whether to enhance the cost of private cars like congestion charging or decreasing the cost of public transit liking travel subsidy is still under discussion. This paper intends to predict the possible effects if the car cost is increased by 10% or the public transportation cost is decreased by 10%, on the basis of the chosen best model. Since the change of cost may directly affect utility of certain mode, the market share of each mode could be affected as a result and this is what we focus on as the main effect cost change may have.

As the data shows, the actual choices of mode are shown in the table below.

Table 5. Actually observed choice

	Walk	Bike	Car	PT	Total
Number of choices	467	690	3229	3800	8186

Market share	5.70%	8.43%	39.45%	46.42%	100.00%

The results of market share under different policy are shown in the table below.

Table 6. Results of policy change

Market share	Walk	Bike	Car	PT	Total
Originally predicted	5.65%	8.44%	39.44%	46.46%	100.00%
Car cost increase by 10%	5.66%	8.52%	39.02%	46.80%	100.00%
PT cost decrease by 10%	5.57%	8.27%	38.86%	47.30%	100.00%

Seen from the tables above, the originally predicted market share without any policy change is slightly overestimated for all the modes except walk. And after increase the car cost by 10%, the market share of car decreases by 0.42%, while PT share increases most by 0.34%, and walk and bike increase slightly. After decrease the public transport cost by 10%, the market share of PT increases by 0.84%, while car share decreased most by 0.58%, and walk and bike increase slightly. These results imply that car and PT share interact more than the other two modes, for the possible reason that they in total count for the majority of the market share overall.

6. Conclusions

This paper estimates four models of three kinds on the basis of the travel data collected in Stockholm City. MNL is the simplest model, while nested model takes into consideration possible correlation between alternatives and mixed model allows coefficient to vary among the population. The latter two models can complement shortcoming of the MNL to some extent, but in the sense of statistical estimation they may not be better than MNL to represent the actual situation. In the case study, by comparison, the MNL is chosen to be best estimated model. Of course, there could be some omissions in the estimation process of the latter two kinds of models as the barrier for a better estimation. In terms of the policy analysis, market shares of car and PT interact obviously given a cost change. And 10% decrease of PT cost seems to have a better promoting effect of PT share increase than 10% increase of car cost.

Nested logit model and mixed logit model can include more information than the MNL. The reasons why these two models don't show good statistical feature in this paper can be various. The inclusion of variables and choice of mixed variables could easily affect estimation result. Even if like this, nested logit and mixed logit deserve further application.

References

Qin, H., Guan, H., Yin, H., et al., 2008. A Study of the effect of parking price on the mode of inhabitant trip behavior with the cars, public transit and taxi in Beijing as an example. China Civil Engineering Journal, 41(8), 93-98.

Zong F., Juan Z., 2007. Activity-based travel mode choice model and analysis on traffic demand management. Journal of Jilin University Engineering and Technology Edition, 37(1),48-53.

Chen Y., Song Y., Zhang Y., et al., 2011. Impact of Land Use Development on Travel Mode Choice: A Case Study in Shenzhen. Urban Transport of China, 5, 80-85.

Zhang N., Dai J., Zhang X., et al., 2012. Walking Affect Area of Rail Transit Station Based on Multinomial Logit Model. Urban Mass Transit, 15(5), 46-49.

Zhang Z., Guan H., Qin H., Xue Y., 2014. Feasibility study of car owner transferring to bus use. Transportation Standardization, 42(17), 5-8.

Liu Z., Zhou X., 2006. Application on nested Logit mode of traffic mode choice . Journal of Shanghai Maritime University, 27(3), 66-70.

Zhang R., Yao E., Yang Y., 2014. Multimodal Urban Traffic Assignment Analysis. Journal of Transportation Systems Engineering and Information Technology, 6, 107-112.

Luo Q., Sun B., Wu W., et al., 2010. Predicting traffic mode split under congestion pricing based on mixed logit model. Journal of Jilin University(Engineering and Technology Edition),40(5), 1230-1234.

Wu L., Lu H., Wang Y., Zhang M., 2014. Research on Intercity Travel Mode Dynamic Choice Behavior with Introduced Loyalty Variable. Journal of Highway and Transportation Research and Development, 31(11), 123-129.

Kenneth E. Train, 2009. Discrete choice methods with simulation. Cambridge university press.

Lu H., 2006. Theory and Method in Transportation Planning. Tsinghua University Press.