Analysis of Mode Choice Behaviours based on Latent Class Models

Back Jin Lee, Hiroshima University
Akimasa Fujiwara, Hiroshima University
Junyi Zhang, Hiroshima University
Yoriyasu Sugie, Hiroshima University

Conference paper
Session V: SP ISSUES

Moving through nets:
The physical and social dimensions of travel
10th International Conference on Travel Behaviour Research
Lucerne, 10-15. August 2003

1. Introduction

In the context of the analysis of travel choice behaviours, the problem of heterogeneity caused by not being able to include particular important explanatory variables in the travel choice behaviours models, has received a great deal of attention. One of the common approaches considering the unobserved heterogeneity is the mixed logit model (MXL) that allows model parameters to vary over individuals. Some parameters of the MXL models are assumed as random values with distribution. While these procedures explicitly incorporate and account for heterogeneity, they are not well suited to explaining the sources of heterogeneity. In many cases these sources relate to the characteristics of individual (Boxall and Adamowicz, 1999).

Alternative approach to consider the heterogeneity is the latent class (LC) model that is commonly adopted in marketing analysis. In this approach, each latent class consists of a number of individuals that are assumed to be homogeneous with respect to their preferences for alternatives as well as their sensitivity to the variations of alternative variables. Latent classes, however, differ in both preference and responsiveness, meaning that the taste parameters differ between latent classes. Swait and Sweeney (2000) also point out that the latent classes can be characterized by variance differences, meaning that the scale parameters differ between latent classes. One restriction of their approach is that either the taste parameters or scale parameters of latent classes should be constrained to equal to one, meaning that they cannot be estimated simultaneously. Each of parameter restrictions represents different behavioural assumptions concerning taste heterogeneity and error term variance within latent classes. While some sources of individual heterogeneity can be explained in the LC model in which the heterogeneity is explained based on the observed factors such as demographic variables, the unobserved factors affecting the heterogeneity cannot be considered in the model.

An approach to combine the LC and the MXL models is worthy to simultaneously incorporate the across-latent class heterogeneity by the LC model and the intra-latent class heterogeneity by the MXL approach into a model. In this paper, the combined model is called as LC-MXL model. Allowing the inter- as well as intra-latent class heterogeneity can be expected to have implications for the ability of the models to fit and the estimated number of workable latent classes.

In this paper, we analyse the mode choice behaviours under traffic information. Stated preference survey was carried out with respect to the choice of the preferred mode: new transit system (NTS) and private car. The main purpose of the survey was to measure the responsiveness for the supplied traffic information about delay time travel information. To collect the indicator data that are measurable variables to infer information about latent variables, we asked to respondents to rate their satisfaction levels with 5 scales for the 14 statements about the alternatives; NTS and car. The subjective rating data for the perceived satisfaction levels will be applied to consider the heterogeneity in the LC model.

The aims of this paper are 1) to compare the MXL model with the LC model, 2) to suggest the LC-MXL model considering the inter- and intra-latent class heterogeneity, 3) to clarify the sources of the individual heterogeneity in the models, and 4) finally to evaluate the effect of supplying traffic information in travellers' mode choice behaviours.

The objective of the comparison of the MXL and LCM models is not to decide a more preferred model but to understand the relative advantages of both models. For instance, it is well known that the LCM model is more cost effective and easy to model specification while the MXL model allows more flexible forms of discrete choice models. It is expected from this paper that (1) more realistic models considering the individual heterogeneity can be developed by adopting the individuals' attitudes and perceptions for alternatives, (2) the comparison of the MXL and LCM models can support to understand advantages/disadvantages of both models, (3) the sources of the individual heterogeneity can be clarified, and (4) the LC-MXL model can enhance the explanatory power of the LC and the MXL models by allowing the inter- and intra-latent class heterogeneity.

This paper is constituted with the following chapters. In chapter 2, the modelling framework of the LC and the MXL models are briefly explained, and also the formulation of LC-MXL model is described in this chapter. In chapter 3, the empirical data and survey method of this paper is described. In chapter 4, the results of parameter estimates are explained. The final conclusions and future works will be summarized in the chapter 5.

2. Model Formulation

We start with the usual assumptions of random utility theory, that when facing a choice situation, individuals assign random utilities to each alternative considered and then choose the alternatives having the highest derived utility. The individual's derived utility can be decomposed into a deterministic component and a random component.

Let us define the random utility assigned to alternative *j* by individual *n* is

$$U_n(j) = \theta_{jn} + \beta_n X_{jn} + \varepsilon_{jn} \tag{1}$$

where θ_{jn} is the intrinsic utility of alternative *j* for individual *n*, β_n is a vector of parameters estimated for individual *n*, X_{jn} is a vector of the attributes of alternative *j* for individual *n*, and ε_{jn} is a random error.

If the random error term is assumed to be independently and identically distributed (iid) extreme value distributions, the multinomial logit (MNL) model can be applied for the conditional probability of choosing alternative *j*. This model inherently assumes the homogeneous population with respect to individuals' preferences for alternatives and their sensitivity to attributes of alternatives. The conditional probability therefore is given in equation (2).

$$P_n(j) = \frac{\exp(\theta_j + \beta X_{jn})}{\sum_{j'=1}^{J} \exp(\theta_{j'} + \beta X_{j'n})} \tag{2}$$

The most well known disadvantages with the MNL model are the parameters of the MNL are fixed in the population (the assumption of homogenous population, which is ignoring taste variations of individuals) and its inherent assumption of independence from irrelevant alternative (IIA). Various alternative formulations have been suggested to relieve the inherent assumption of the MNL model. One of the most significant of these extensions is the mixed logit (MXL) model that allows model parameters to vary over individuals. Some parameters of the MXL models are assumed as random values with distribution. Another approach is the latent class (LC) model. In this approach, each latent class consists of a number of individuals that are assumed to be homogeneous with respect to their preferences for alternatives as well as their sensitivity to alternative variables. In this chapter, we will briefly explain about the MXL model and LC model. Furthermore, the LC-MXL model will be described.

2.1 Latent Class (LC) Model

We begin by assuming that there exist *S* latent classes in the heterogeneous population under investigation (*s=1, 2, …, S*). Each latent class consists of a number of populations that are assumed to be homogeneous with respect to their preferences for travel modes as well as their sensitivity to alternative variables. Latent classes, however, differ in both preferences and sensitivity, meaning that the taste parameters differ between latent classes.

Let us define $P_{n/s}(j)$ as the probability that individual *n* belonging to latent class *s* chooses alternative *j*. Given the traditional random utility framework and the assumption of iid extreme value distributions for the random error term, this conditional probability can be written as

$$P_{n/s}(j) = \frac{\exp(\theta_{js} + \beta_s X_{jn})}{\sum_{j'=1}^{J} \exp(\theta_{j's} + \beta_s X_{j'n})} \tag{3}$$

where θ_{js} is the intrinsic utility of alternative *j* for individual *n* in segment *s*, and β_s is the parameter vector estimated for individual *n* in segment *s*.

The probability that an individual *n* belongs to latent class s, P_{ns} is assumed to relate with the individual's latent attitudes and perceptions. This probability can be represented as

$$P_{ns} = \frac{\exp(\alpha_s + \gamma_s Z_n)}{\sum_{s'=1}^{S} \exp(\alpha_{s'} + \gamma_{s'} Z_n)} \tag{4}$$

where α_s is the intercept and γ_s are unknown latent class parameters to be estimated and denote the contribution of the latent attitudes and perceptions, Z_n, to the probability of latent class membership. Boxall and Adamowicz (1999) assumed that Z_n is a vector of both the psychographics constructs and socio-economic characteristics. In this paper, we assume that Z_n is a vector of the psychographics constructs only for the simplification of the model. The following constraints must be met in equation (4):

$$\sum_{s'}^{S} P_{ns'} = 1\,,\ 0 < P_{ns} \leq 1\,,\ \text{and}\ \alpha_S = \gamma_S = 0 \tag{5}$$

Based on equation (3) and (4), the probability that a randomly selected individual *n* choose alternative *j*, $P_n(j)$ is written in equation (6)

$$P_n(j) = \sum_{s'}^{S} P_{ns'} \cdot P_{n/s'}(j) \tag{6}$$

The likelihood function for individual *n*, L_n is given by

$$L_n = \sum_{s'}^{S} P_{ns'} \cdot \prod_{j'}^{J} P_{n/s'}(j')^{\delta_{j'n}} \tag{7}$$

where δ_{jn} is 1 if individual *n* choose alternative *j*, otherwise 0. Consequently, the sample log-likelihood function is given by

$$LL = \sum_{n=1}^{N} \log(L_n) = \sum_{n=1}^{N} \log\{\sum_{s'}^{S} P_{ns'} \cdot \prod_{j'}^{J} P_{n/s'}(j')^{\delta_{j'n}}\} \tag{8}$$

The likelihood equation (8), however, does not yield an explicit solution for the unknown parameters. Therefore, expectation-maximization (EM) algorithm will be employed to estimate these parameters. The primary advantages of this algorithm are numerical stability, simplicity, and a factorisation of the likelihood function (Mclachlan and Krishnan, 1997). The EM algorithm can be used to find the solution of the optimisation problem.

In the EM algorithm, the estimation of the sample log-likelihood function, equation (8), is treated with the missing data problem. Let us define Z_{ns}, with $Z_{ns} = 1$ if individual *n* belongs to latent class *s* and $Z_{ns} = 0$ otherwise.

If the Z_{ns}'s were known, the complete log-likelihood function, equation (8), can be written as

$$\begin{aligned} LL &= \sum_{n=1}^{N} \log(L_n) = \sum_{n=1}^{N} \log\{\prod_{s'}^{S} P_{ns'} \prod_{j'}^{J} P_{n/s'}(j')^{\delta_{j'n}}\}^{z_{s'n}} \\ &= \sum_{n=1}^{N}\sum_{s'}^{S} z_{ns'} \log(P_{ns'}) + \sum_{n=1}^{N}\sum_{s'}^{S}\sum_{j'}^{J} z_{ns'}\delta_{j'n} \log\{P_{n/s'}(j')\} \end{aligned} \tag{9}$$

The EM algorithm comprises two steps. In the expectation step, the expected probability of the latent is computed using a set of initial values for the model parameters. The posterior probability that a individual *n* belongs to latent class *s*, $\hat{Z}_{ns}$, can be obtained as

$$\hat{z}_{ns} = \frac{P_{ns} \cdot \prod_{j'}^{J} P_{n/s}(j')^{\delta_{j'n}}}{\sum_{s'=1}^{S} P_{ns'} \cdot \prod_{j'}^{J} P_{n/s'}(j')^{\delta_{j'n}}} \tag{10}$$

The E step of the EM algorithm thus amount to replacing the unobservable Z_{ns} in the complete log-likelihood function by their current expectation, $\hat{Z}_{ns}$. The last part of the EM algorithm is the M step, which is to maximize equation (9). The maximum likelihood parameter estimates then replaces the initial parameters to develop new expected probability ($\hat{z}_{ns}$), and these steps are repeated until the parameter estimates achieve convergence. During the iteration, the likelihood value increases monotonously (Mclachlan and Krishnan, 1997).

One of the major difficulties in applying the latent class approach is determining the “correct” number of latent classes. Typically, this decision is based on information criteria such as the Bayesian Information Criterion (BIC) or the Akaike Information Criterion (AIC). In this paper, the both criteria of BIC and AIC will be employed to decide the number of latent classes

$$BIC = -2 \cdot \{LL(\beta)\} + p \cdot \log(N) \tag{11}$$

$$AIC = -2 \cdot \{LL(\beta) - p\} \tag{12}$$

where $LL(\beta)$ is the log-likelihood value at convergence, p is a number of parameters, and N is the number of samples.

2.2 Mixed Logit Model

In the form of mixed logit model, the individual heterogeneity can be considered by varying the parameters in the population, random parameters with the mean and variance. The individual heterogeneity can be modelled by assuming that the parameters of equation (1) follow a multivariate normal distribution across the population of individual.

Let us assume that the parameters, β_n, follow a certain distribution with the mean, β_{mean}, and the standard deviation, $\beta_{std.}$, then the equation (1) can be rewritten as

$$\begin{aligned} U_n(j) &= \theta_{jn} + f(\beta_n \setminus \beta_{mean}, \beta_{std.}) X_{jn} + \varepsilon_{jn} \\ &= \theta_{jn} + \beta_{mean} X_{jn} + \beta_{std.} X_{jn} + \varepsilon_{jn} \\ &= \theta_{jn} + \beta_{mean} X_{jn} + \eta_{jn} + \varepsilon_{jn} \end{aligned} \tag{13}$$

where $\beta_{mean} X_{jn}$ is non-stochastic, linear-in-parameters par that depends on observed data. The parameter β_{mean} is fixed in the population, $\beta_{std.} X_{jn} (= \eta_{jn})$ is recognized as the random error term that induce individual heterogeneity. The parameters η_{jn} vary in the population with the density $f(\eta_{jn} \setminus \lambda)$, where λ are the true parameters of the distribution. Different patterns of individual heterogeneity can be obtained through different specifications of $f(\eta_{jn} \setminus \lambda)$ such as normal, uniform, triangular and lognormal distributions. ε_{jn} is an unobserved random term and independently, identically distributed (iid) over alternatives and individuals. If we know η_{jn}, the conditional probability that an individual *n* chooses alternative *j* is standard logit.

$$P_n(j/\eta_{jn}) = \frac{\exp(\theta_{jn} + \beta_{mean} X_{jn} + \eta_{jn})}{\sum_{j'=1}^{J} \exp(\theta_{j'n} + \beta_{mean} X_{j'n} + \eta_{j'n})} \tag{14}$$

In practical, we do not know the individual's tastes. Therefore, we need to calculate the unconditional probability obtained by integrating equation (14) over all possible values of η_{nj}.

$$\begin{aligned} P_n(j \mid \lambda) &= \int P_n(j \mid \eta_{jn}) f(\eta_{jn} \setminus \lambda) \\ &= \int \frac{\exp(\theta_{jn} + \beta_{mean} X_{jn} + \eta_{jn})}{\sum_{j'=1}^{J} \exp(\theta_{j'n} + \beta_{mean} X_{j'n} + \eta_{j'n})} f(\eta_{jn} \setminus \lambda) d\eta_{jn} \end{aligned} \tag{15}$$

Because the choice probability is a mixture of logit with *f* function as the mixing distribution, we call equation (15) is the MXL model. The integral in equation (15) cannot be estimated analytically because it does not have a closed form. Therefore, the choice probability is approximated through simulation. The brief explanation of the estimation is as follows; In the first step, the unobserved parameter η_{nj} is randomly drawn from the distribution that is assumed as normal or log normal with initial mean and variance. In the second step, using the drawn η_{nj}, the standard logit formula is calculated. Repeat the first and second steps *R* times, thus obtaining *r* values for the likelihood function l_r. In the third step, compute the average with the following equation (16)

$$SP_r = \frac{1}{R}\left(\sum_{r}^{R} l_r\right) \tag{16}$$

where SP_r is the simulated likelihood value and unbiased estimator.

2.3 LC-MXL models

The combined approach of the LC and the MXL models is worthy to simultaneously incorporate the across-latent class heterogeneity by the LC model and the intra-latent class heterogeneity by the MXL model. In this paper, the combined model is called as LC-MXL model. In this paper, we apply a sequential approach to the combined model. In the first step, the LC model will be estimated to decide the optimal number of latent classes and to incorporate the inter-latent class heterogeneity. Based on the estimation results of the LC model, in the second step, the MXL model will be simultaneously estimated for all latent classes. Ideally the combined approach is necessary to simultaneously estimate the LC and the MXL models. This is left for future works of this paper.

The formulation of the combined model is simply the extension of equation (13). Therefore, equation (13) can be rewritten as

$$\begin{aligned} U_n(j) &= \theta_{js} + f(\beta_s \backslash \beta_{s\_mean}, \beta_{s\_std.}) X_{jn} + \varepsilon_{jn} \\ &= \theta_{js} + \beta_{s\_mean} X_{jn} + \beta_{s\_std.} X_{jn} + \varepsilon_{jn} \\ &= \theta_{js} + \beta_{s\_mean} X_{jn} + \eta_{s\_jn} + \varepsilon_{jn} \end{aligned} \tag{17}$$

where $f(\beta_s \backslash \beta_{s\_mean}, \beta_{s\_std.})$ means that the parameters of each latent class follow a certain distribution with the mean, β_{s_mean}, and the standard deviation, $\beta_{s\_std.}$. Therefore, the β_{s_mean} and $\beta_{s\_std.} X_{nj}$ represent the inter- and intra-latent class heterogeneity, respectively.

Although the sequential approach for the estimation of the combined model is applied in this paper, we expect that the simultaneous estimation also can be easily implemented by developing the existing LC or MXL approaches.

3. Empirical Study

In this research, we analyze individuals' mode choice behaviors of car and new transit systems (NTS), when delay time traffic information for car use is supplied for car users. In addition, we use individuals' subjective attitudes and perceptions of alternatives (car and NTS) for analyzing the individual heterogeneity in mode choice behaviors.

The data used in this study were collected through a questionnaire survey. The survey area is located at the northwestern residential area along NTS in Hiroshima. Stated preference survey is carried out to investigate the car users' intentions to choose NTS when delay time traffic information is supplied for the car users. In addition, respondents' perceived satisfaction ratings for car, NTS, and access bus to NTS stations were collected. The respondents rated their perceived satisfaction on a set of fourteen descriptive attributes shown in Table 1. Five linguistic levels were used for satisfaction ratings: very satisfied, satisfied, normal, dissatisfied, and very dissatisfied.

Table 1 Subjective Rating Item

Alternatives	Rating Item
Car	Var1: Travel time
	Var2: Congestion level
	Var3: Traffic information supplied by radio
	Var4: Traffic information supplied by TV
Access bus	Var5: Ordinary service level
	Var6: Service level when bad weather (raining, snow)
	Var7: Service frequency
	Var8: Travel time
	Var9: Fare
NTS	Var10: Service level when bad weather (raining, snow)
	Var11: Service frequency
	Var12: Service schedule
	Var13: Travel time
	Var14: Fare

In SP experiments, SP factors were 1) whether delay time information is supplied or not, 2) delay time of car mode and 3) travel time of NTS. Four types of profiles were set up based on the levels of SP factors, shown in Table 2; SP1 scenario represents ordinary situations in where delay time traffic information is not supplied; SP2 scenario represents that no delay time traffic information is supplied for car users; SP3 scenario represents when traffic information of 10 or 20 minutes delay time due to light congestions is supplied; SP4 scenario when traffic information of 30 or 40 minutes delay time due to heavy congestions is supplied.

Table 2 SP scenarios

Card	SP1	SP2	SP3	SP4
Traffic information	Not supplied	Supplied	Supplied	Supplied
Delay time for car	-	0	10,20	30,40
Travel time of NTS	20	20	20	20

Travel time of NTS is fixed to 20 minutes in all SP scenarios. The total travel time of car is respondent's perceived travel time plus supplied delay time information.

3.1 Factor analysis

To incorporate the unobserved heterogeneity of individuals into the LC model, the factor analysis is implemented by using the respondents' perceived satisfaction rating data representing individuals' attitudes and perceptions for alternatives. The factor analysis provides estimates that enter the membership likelihood function. The observed satisfaction rating data from the 13 statements were analysed using principal component factor analysis with varimax rotation. The estimation results of the factor analysis are shown in Table 3. Totally, five principal components for respondents' satisfaction rating data are identified from the factor analysis. These components can be defined based on magnitudes of the factor loadings.

- *The first principal factor*: Satisfaction levels for service performance of NTS
- *The second principal factor*: Satisfaction levels for service performance of access bus
- *The third principal factor*: Satisfaction levels for traffic information for car
- *The forth principal factor*: Satisfaction levels for fare of NTS and BUS
- *The fifth principal factor*: Satisfaction levels for service performance of car

Factor scores for the five principal factors were then calculated using a regression model for each individual, and consequently these factor scores will be applied as explanatory variables, Z_n, to estimate the probability that an individual *n* belongs to latent class s, P_{ns} is assumed to relate with the individual's latent attitudes and perceptions. Therefore, five explanatory variables and an intercept will be included in the Z_n vector.

Table 3 Factor analyses of attitudinal statements reflecting satisfaction levels for car and NTS

Variables	Factor loadings				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
VAR1	0.088	0.115	0.101	0.106	0.855
VAR2	0.099	0.123	0.104	0.089	0.860
VAR3	0.061	0.127	0.870	0.118	0.127
VAR4	0.104	0.007	0.884	0.018	0.081
VAR5	0.093	0.762	0.088	-0.095	0.065
VAR6	0.027	0.806	0.008	0.157	0.037
VAR7	0.141	0.633	0.108	0.322	0.003
VAR8	0.180	0.656	0.006	0.008	0.223
VAR9	-0.008	0.189	0.065	0.856	0.087
VAR10	0.737	0.124	-0.015	-0.013	0.121
VAR11	0.754	0.235	-0.034	0.148	0.001
VAR12	0.742	-0.013	0.145	0.197	0.062
VAR13	0.750	0.082	0.122	-0.001	0.056
VAR14	0.257	0.014	0.065	0.793	0.119

4. Mode Choice Behaviour Models Considering Individual Heterogeneity

Table 4 Definitions of Explanatory Variables

TT	Travel time of car and NTS (Generic)
DT	Delay time traffic information (Car specific)
DUI	Dummy of traffic information (Car specific)
	If travel information is supplied 1
	Otherwise 0
CC	Constant (Car specific)

Table 4 shows the explanatory variables that are used for the model choice behaviour models. To analyse the effect of delay time traffic information to mode choice behaviours of car users, we developed mode choice behaviour models for the following three cases.

- *Case 1 (SP1)*: Delay time traffic information is not supplied for car users. This presents ordinary conditions.
- *Case 2 (SP1+SP2)*: 0 minutes delay time traffic information is supplied.
- *Case 3 (SP2+SP3+SP4)*: Delay time traffic information is supplied

Four types of mode choice behaviour models are developed for each case.

- *Model 1*-MNL models: No heterogeneity
- *Model 2*-MXL models: Intra heterogeneity
- *Model 3*-LC models: Inter-latent class heterogeneity
- *Model 4*-LC-MXL models: Intra- and Inter-latent class heterogeneity

The model 1 is developed as the “base model”, which ignore the individuals’ heterogeneity. The model 2 is the case of allowing the heterogeneity for the model 1. The model 3 is the case of allowing the inter-latent class heterogeneity, but does not allow the intra-latent class heterogeneity. The model 4 is the case of inter- as well as intra-latent class heterogeneity.

Table 5 Comparison of AIC and BIC

Num. of latent classes	Num. of parameters	Log-likelihood at convergence	AIC	BIC
1	2	-198.363	400.726	408.140
2	10 (4+6)	-184.939	389.878.	426.949
3	18 (6+12)	-148.177	332.354	399.082
4	26 (8+18)	No significant results are obtained		

Note) the sample size is 301 individuals

Table 6 Parameter estimates on the latent class membership variables of the three LC model for the case 1 (t-statistics)

Variables	LC1	LC2	LC3
Intercept	1.957 (4.329)	-5.101 (-3.014)	0
Principal factor 1	3.011 (4.867)	15.262 (5.771)	0
Principal factor 2	3.632 (4.652)	2.134 (2.014)	0
Principal factor 3	3.998 (4.784)	15.465 (6.008)	0
Principal factor 4	1.612 (3.911)	3.036 (4.226)	0
Principal factor 5	2.067 (4.990)	4.174 (4.242)	0

Note: Latent class 3 is treated as the "base" class

Especially, for developing Model 3 and 4, that consider the inter-latent class heterogeneity, we firstly develop the LC model for the case 1, which is assumed to represent the individuals' heterogeneity for mode choice behaviours in ordinary road conditions. With the estimation results of this LC model, it is known that an individual n belongs to latent class s. In this stage, the factor scores for the five principal components and an intercept are applied as explanatory variables to estimate the membership probability in the LC model. Based on the classification results of the LC model for case 1, we develop the model 3 and 4 for the case 2 and 3 to ex-

Table 7 Parameter Estimates for the Case 1 (t-statistics)

	MNL	MXL	3-LC model			3-LC-MXL		
			LC1	LC2	LC3	LC1	LC2	LC3
TT	-0.023	-0.018	-0.227	-0.125	-0.159	-0.198	-0.108	-0.169
	(-2.098)	(-0.295)	(-3.969)	(-2.965)	(-3.956)	(-3.711)	(-3.570)	(-2.811)
CC	0.898	1.117	-1.103	2.293	1.689	-0.911	2.079	1.762
	(3.728)	(2.362)	(-1.960)	(3.270)	(3.239)	(-1.498)	(4.264)	(2.782)
Std. dev. TT		0.213				0.002	0.020	0.028
		(0.539)				(0.186)	(0.408)	(0.469)
N.O.S	301	301	103	93	105	103	93	105
LL (0)	-208.637	-208.637		-208.637			-208.637	
LL (β)	-198.363	197.960		-148.177			-151.832	
Rho (ρ)	0.049	0.051		0.290			0.272	
A-rho (ρ)	0.040	0.037		0.204			0.215	
N.O.P	2	3		18			12	

Note: N.O.S and N.O.P are the number of sample size parameters, respectively
A-rho (ρ) is the adjusted rho-square of degree of freedom

amine the different responsiveness between latent classes. Specifically, the model 3 and 4 for the case 2 and 3 are developed with the prior segmentation based on the case 1.

In estimating the LC model for the case of 1, the number of latent classes, from 1, 2, 3, and 4, are attempted and the comparison results of AIC and BIC values of these LC models are summarized in Table 5. In the case of when the number of latent classes is four, no significant estimation results are obtained; we consider that the reason is because of the number of sample size. The comparison results of BIC and AIC reveals that the optimal number of latent classes is four since both the lowest BIC and AIC values are obtained in that case. Therefore, we use the estimation results of the 3-LC model for the further analysis. The results of parameter estimates for the case 1 and the 3-LC model are shown in Table 6 and 7. From Table 6, we note that latent class 2 are constituted with individuals having the highest satisfaction levels for the service performance of NTS, traffic information for car, and service performance of car, but they have the lowest satisfaction levels for the service performance of access bus. However, it is possible to exist the intra-latent class heterogeneity in the latent class 2, since the individuals of the latent class 2 may have the quite different satisfaction levels for NTS and car modes. We can also define that the latent class 1 are constituted with individuals having the lowest satisfaction levels for the fare of NTS and access bus.

The MXL model, in Table 7, is developed by assuming that the TT parameters of individuals are normally distributed. However, the MXL model considering the individuals heterogeneity does not improve the MNL model. The standard deviation of TT is not significant, and this represents that the heterogeneity can be ignored. On the other hand, the 3-LC model significantly improves the explanatory power of the MNL model by incorporating the inter-latent class heterogeneity. From this fact, it is known that the MXL model cannot completely capture the individual (unobserved) heterogeneity. The estimation results of the 3-LC-MXL model shows that there is no significant intra-latent class heterogeneity, and these results are similar with those of the 3-LC models. Note that the 3-LC-MXL models are estimated based on the prior segmentation of 3-LC model, and the further models of the 3-LC and 3-LC-MXL models for the case 2 and 3 are also estimated with the same way.

To estimate the different responsiveness of individuals, we estimate the models for the case 2 and 3. From Table 8 we can note that the MXL model incorporating the individual heterogeneity still not improve the MNL model. However, we found that the individual heterogeneity exists in the DUI representing the responsiveness whether 0 minutes delay time traffic information for car provided or not. The estimation results of the 3-LC model represent that the individuals of latent class 1 have the highest responsiveness for the DUI, and this model significantly improve the MNL model by considering the inter-latent class heterogeneity. From the estimation results of the 3-LC-MXL model considering both the inter- and intra-latent class heterogeneity, it is known that there exists the intra-latent class heterogeneity. Specifically, the intra-latent class heterogeneity for the DUI exist in all the latent classes, and for TT exists in the latent class 2. In addition, the 3-LC-MXL model little improve the explanatory power of the 3-LC model.

From Table 9, it is noted that the MXL model is significantly better than the MNL model at the 5% level of significance ($\kappa^2_{0.05}$=5.99< -2*(526.276-530.518)=8.484). The estimation results of the 3-LC model show that individuals of latent class 1 are the most sensitivity for the delay time traffic information of car. In addition, the 3-LC model is significantly better than

Table 8 Parameter Estimates for the Case 2 (t-statistics)

	MNL	MXL	3-LC model			3-LC-MXL		
			LC1	LC2	LC3	LC1	LC2	LC3
TT	-0.029	-0.001	-0.163	-0.092	-0.113	-0.201	-0.249	-0.171
	(-2.782)	(-0.019)	(-4.940)	(-3.649)	(-5.530)	(-2.999)	(-1.881)	(-3.192)
DUI	0.523	4.051	0.834	0.726	0.546	2.401	1.213	0.847
	(2.959)	(1.413)	(1.881)	(2.155)	(1.758)	(1.160)	(1.516)	(1.518)
CC	0.885	1.301	-0.570	2.293	1.133	-0.930	-5.435	1.775
	(4.561)	(3.225)	(-1.355)	(3.270)	(3.480)	(-1.445)	(-1.798)	(3.020)
Std. dev. TT		0.213				0.017	0.447	0.035
		(0.539)				(0.117)	(1.421)	(0.793)
Std. dev. DUI		6.587				2.644	1.168	2.852
		(1.495)				(1.269)	(1.549)	(1.664)
N.O.S	602	602	206	186	210	206	186	210
LL (0)	-417.275	-417.275		-417.275			-417.275	
LL (β)	-372.924	-371.322		-301.257			-297.197	
Rho (ρ)	0.106	0.110		0.278			0.287	
A-rho (ρ)	0.099	0.098		0.256			0.275	
N.O.P	3	5		9			15	

the both MNL and MXL models by considering the inter-latent class heterogeneity. From the results of parameter estimates of the 3-LC-MXL model, this model little improve the 3-LC model by considering the intra-latent class heterogeneity. Especially, the intra-latent class heterogeneity exists for the latent class 2 in terms of TT and DT variables.

Table 9 Parameter Estimates for the Case 3 (t-statistics)

	MNL	MXL	3-LC model			3-LC-MXL		
			LC1	LC2	LC3	LC1	LC2	LC3
TT	-0.010	-0.025	-0.070	-0.050	-0.063	-0.078	-0.144	-0.188
	(-1.683)	(-1.537)	(-3.947)	(-2.807)	(-4.174)	(-2.999)	(-1.611)	(-1.791)
DT	-0.068	-0.098	-0.101	-0.067	-0.059	-0.112	-0.186	-0.112
	(-11.116)	(-5.807)	(-8.652)	(-6.683)	(-6.202)	(-0.264)	(-1.634)	(-2.099)
CC	1.122	1.702	0.996	1.796	1.077	1.023	4.543	2.521
	(7.962)	(4.512)	(3.091)	(4.974)	(3.837)	(0.762)	(1.899)	(2.052)
Std. dev. TT		0.110				0.022	0.290	0.226
		(2.386)				(0.022)	(1.379)	(1.240)
Std. dev. DT		0.019				-0.030	0.085	0.009
		(0.849)				(0.068)	(1.223)	(0.515)
N.O.S	903	903	309	279	315	309	279	315
LL (0)	-625.912	-625.912		-625.912			-625.912	
LL (β)	-530.518	-526.276		-486.445			-480.379	
Rho (ρ)	0.152	0.160		0.223			0.233	
A-rho (ρ)	0.148	0.151		0.208			0.209	
N.O.P	3	5		9			15	

By comparing the outputs of the case 1 and the case 3, it can be emphasized that the MXL model can more effectively capture the individuals' heterogeneity when the traffic information is supplied. Specifically, employing the MXL model can capture the individual heterogeneity of the responsiveness for the traffic information. In the MXL model, however, it is difficult to represent or analyse the reasons why the individual heterogeneity arise. On the other hand, the LC model using the factor analysis to capture the unobserved (latent) heterogeneity is able to represent the reasons.

Moreover, The estimation results of the LC-MXL models, that allow the inter- as well as intra-latent class heterogeneity, show higher adjusted rho-square comparing those of the LC models, that allow the inter-latent class heterogeneity only; 0.204→0.215 (case 1), 0.256→0.275 (case 2), and 0.208→0.209 (case 3). This result shows that incorporating inter- and intra-latent class heterogeneity can enhance the explanatory power of mode choice behaviour models as well as the representation ability of the individual heterogeneity.

5. Conclusions

The aims of this paper are to compare the MXL model with the LC model, to suggest and verify the LC-MXL model considering the inter- and intra-latent class heterogeneity, and to analyse the sources of the individual heterogeneity in terms of mode choice behaviours.

Based on the estimated models considering the individual heterogeneity, some results of this paper can be summarized as follows:

(1) The sources of the individual heterogeneity can be represented using the latent variables such as individuals' attitudes and perceptions for the corresponding alternatives. In this paper, the factor analysis is implemented to induce the principal latent factors from the individuals' subjective rating data.

(2) The MXL model can more effectively capture the individuals' heterogeneity for the responsiveness for the supplied traffic information. In the MXL model, however, it is difficult to represent or analyse the reasons why the individual heterogeneity arise. On the other hand, the LC model using the factor analysis to capture the unobserved (latent) heterogeneity is able to represent the reasons.

(3) The estimation results of the LC-MXL models, that allow the inter- as well as intra-latent class heterogeneity, shows that incorporating inter- and intra-latent class heterogeneity can enhance the explanatory power of mode choice behaviour models as well as the representation ability of the individual heterogeneity

6. References

1) Ben-Akiva, M., McFadden, D., Train, K. and Walker, J. (2001) Hybrid choice models: progress and challenges, working paper, MIT.

2) Boxall, P.C. and Adamowicz, W.L (1999) Understanding heterogeneous preferences in random utility models: The use of latent class analysis, Staff paper 99-02, Department of Rural Economy, University of Alberta Edmonton, Canada.

3) Back-Jin, L., Fujiwara, A., Sugie, Y. and Namgung, M. (2003): A sequential method for combining random utility model and fuzzy inference model, Journal of Advanced Computational Intelligence and Informatics, (in press).

4) Dempster, A.P., Laird, N.M., and Rubin, D.B. (1997) Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal statistical Society B, Vol.39, pp.1-38.

5) Desarbo, W.S., Kamakura, W.A., and Wedel, M. (2002) Applications of multivariate latent variable models in marketing, Working paper, Pennsylvania State University.

6) Dobson, R. (1975) Towards the Analysis of attitudinal and behavioural responses to transportation system characteristics, Transportation, Vol.4, pp.267-290.

7) Dobson, R. and Tischer, M.L. (1977) Comparative analysis of determinants of modal choice by central business district workers, Transportation Research Record, No.649, pp.7-14.

8) Goulias, K.G. (1999) Longitudinal analysis of activity and travel pattern dynamics using generalized mixed Markov latent class models, Transportation Research Part B, No.33, pp.535-557.

9) Gupta, S. and Chintagunta, P.K. (1994) On using demographic variables to determine segment membership in logit mixture models, Journal of Marketing Research, Vol.31, pp.128-136.

10) Green, W.H. and Hensher, D.A (2002) A latent class model for discrete choice analysis: Contrasts with mixed logit, Working paper, ITS-WP-02-08, Institute of transport studies, The University of Sydney and Monash University, Sydney.

11) Hagennaars, J.A. and McCutcheon, A.L. (2002) Applied latent class analysis, Cambridge University press, Cambridge, UK.

12) Hing-Po, L. and Wendy, L. (2000) A latent class model applied to stated preference data, Application Workshops, The 9th meeting of IATBR 2000, Queensland, Australia.

13) Hutchinson, J., Kamakura, W., and Lynch, J. (2000) Unobserved heterogeneity as an alternative explanation for reversal effects in behavioural research, Journal of consumer research, Vol.27, pp. 324-344.

14) Jain. D., Bass, F.M. and Chen, Y.M. (1990) Estimation of latent class models with heterogeneous choice probabilities-An application to market structuring, Journal of Marketing Research, Vol. 27, pp. 94-101.

15) Kamakura, W.A. and Russell, G.J. (1989) A probabilistic choice model for market segmentation and elasticity structure, Journal of Marketing Research, Vol.26, pp.379-390.

16) Koppleman, F.S. and Lyon, P.K. (1980) Travel choice behaviour: Models of perceptions, feelings, preference, and choice, Transportation Research Record, No.765, pp.26-33.

17) Koppleman, F.S. and Lyon, P.K. (1981) Attitudinal analysis of work/school travel, Transportation Science, Vol.25, No.3, pp.26-33.

18) Loehlin, J.C. (1998) Latent variable models-An introduction to factor, path, and structural analysis, Lawrence Erlbaum Associates, Inc.

19) Louviere, J.J, Hensher, D.A., and Swait, J.D. (2000) Stated Choice Models: Analysis and Application, Cambridge, UK.

20) Lovelock, C.H. (1975) Modelling the modal choice decision process, Transportation, Vol.4, pp.253-265.

21) McFadden, D. (1986) The choice theory approach to market research, Marketing Science, 5:275-297.

22) McFadden, D. (1999) Rationality for economists? Journal of Risk and Uncertainty, Vol.19 (1-3), pp.73-105.

23) McLachlan, G.J., Krishnan, T. (1997) The EM Algorithm and Extensions. John Wiley and Sons, New York, USA.

24) McLachlan, G.j., and Basford, K.E. (1998) Mixture models: Inference and applications to clustering, Marcel Dekker, New York.

25) Morikawa, T. and Sasaki, K. (1998) Discrete choice models with latent variables using subjective data, In Ortuzar, J., Hensher, D. and Jara-Diaz, S. (eds), Travel Behaviour Research: Updating the state of play, Pergamon, pp.435-455.

26) Poulsen, C.S., Brockhoff, P.M.B, and Erichsen, L. (1997) Heterogeneity in consumer preference data-A combined approach, Food Quality and Preference, Vol.8, No.5/6, pp.409-417.

27) Recker, W.W. and Golob, T.F. (1976) Attitudinal modal choice model, Transportation Research, Vol.10, No.5-B, pp.299-310.

28) Swait, J. and Bernardino, A. (2000) Distinguishing taste variation from error structure in discrete choice data, Transportation Research Part B, No.34, pp.1-15.

29) Swait, J. And Sweeney, J.C. (2000) Perceived value and its impact on choice behaviour in a retail setting, Retailing and consumer services, Vol.7, 77-88.

30) Vythoulkas, P.C. and Koutsopoulos, H.N. (2003) Modelling discrete choice behaviour using concepts from fuzzy set theory, approximate reasoning and neural networks, Transportation Research Part C, Vol.11, pp.51-73.

31) Walker, J.L. (2001) Extended discrete choice models: integrated framework, flexible error structures, and latent variables, Ph.D. Dissertation, MIT.

32) Walker, J. and Ben-Akiva, M. (2002) Generalized random utility model, Mathematical Social Sciences, 43, pp.303-343.

33) Walker, J.L. (2000) Extended discrete choice models: Integrated framework, flexible error structures, and latent variables, Doctoral dissertation, Department of Civil and Environmental Engineering, MIT, Massachusetts.

34) Zenor, M.J. and Srivastava, R.k. (1993) Inferring market structure with aggregate data-A latent segment logit approach, Journal of Marketing Research, Vol.30, pp. 369-379.