TRAVEL BEHAVIOR ANALYSIS UNDER UNCERTAINTIES BY USING LATENT CLASS CLUSTERING

Backjin Lee
Ph.D. Candidate
Graduate School for International Development and Cooperation, Hiroshima University
1-5-1, Kagamiyama, Higashi-Hiroshima,
Hiroshima, 739-8529, Japan
Fax: +81-824-24-6922
E-mail: bjlee@hiroshima-u.ac.jp

Akimasa FUJIWARA
Associate Professor
Graduate School for International Development and Cooperation, Hiroshima University
1-5-1, Kagamiyama, Higashi-Hiroshima,
Hiroshima, 739-8529, Japan
Fax: +81-824-24-6921
E-mail: afujiw@hiroshima-u.ac.jp

Yoriyasu SUGIE
Professor
Graduate School for International Development and Cooperation, Hiroshima University
1-5-1, Kagamiyama, Higashi-Hiroshima,
Hiroshima, 739-8529, Japan
Fax: +81-824-24-6919
E-mail: ysugie@hiroshima-u.ac.jp

Moon Namgung
Associate Professor
Department of Civil & Environmental Engineering, Wonkwang University
344-2 Shinyongdong, Iksanshi, Chollabukdo
570-749 Korea
Fax: +82-61-850-6722
E-mail: ngmoon@wonnms.wonkwang.ac.kr

Abstract: The uncertainties of travel choice behaviors can be divided into two different types, randomness and vagueness. Transportation researchers usually assume only one uncertainty for travel choice behavior modeling, and the assumption is based on that travelers of their sample present a homogeneous population. However, traveler perceptions may differ at the individual level of familiarity with the network attributes. Therefore, it is important to consider the heterogeneity of traveler perceptions in sample data and to employ a different modeling methodology by their uncertainty types. In this paper, a latent class clustering approach is applied for considering the heterogeneity of traveler perceptions in sample data. In addition, random utility models and fuzzy reasoning models considering the heterogeneity are developed to find an appropriate model by the types of uncertainty. All results of the paper emphasize the necessity of a combined model, which can consider the randomness and the vagueness uncertainty simultaneously.

Key Words: Latent class clustering, EM algorithm, Fuzzy reasoning model and Uncertainty

1. INTRODUCTION

Treatment of uncertainties is critical for modeling travel choice behaviors. The uncertainties can be divided into two different types (Lin and George Lee, 1996): One is randomness due to the non-deterministic nature of travel choice behavior problems. Random utility models have been employed to deal with the uncertainty, and probability distribution is applicable for measuring the randomness. The other is vagueness due to the poor knowledge and the lack of familiarity with network attributes (Lotan and Koutshopoulos, 1993[a]). Fuzzy reasoning models have been used to the vagueness of uncertainty, and possibility distribution is appropriate for measuring the vagueness. Transportation researchers usually assume only one uncertainty for travel choice behavior modeling, and the assumption is based on that travelers of their sample present a homogeneous population.

Traveler perceptions, however, may differ at the individual level of familiarity with the network attributes. Lotan and Koutshopoulos (1993[b]) instanced that a traveler familiar with the network attributes is able to derive a distribution of travel times. Therefore, probability measures can be used to model the perceptions of the very familiar traveler. On the other hand uncertainty due to vagueness is mainly related to cases in which a traveler is not familiar with the network attributes and therefore has very little idea about the actual characteristics of that network. Fuzzy reasoning models can be used in this case to model traveler perceptions. Therefore, it is important to consider the heterogeneity of traveler perceptions in sample data and to employ a different modeling methodology by their uncertainty types.

In this paper, route choice behavior models are estimated with the assumption that drivers are mainly influenced by their perceived levels of the travel time on each alternative route. A latent class clustering approach is applied for considering the heterogeneity of driver perceptions in sample data, that is, the heterogeneous population is classified into finite homogeneous subpopulations called as latent classes in the paper. After classifying the sample data into latent classes, random utility models and fuzzy reasoning models considering the heterogeneity are developed to find an appropriate model by the types of uncertainty. In addition, the final objective of the paper is to support the necessity of a combined model considering the randomness and the vagueness uncertainty simultaneously.

The paper is constituted as follows. In chapter 2, the mathematical framework of latent class clustering is introduced. In chapter 3, the characteristics of random utility model and fuzzy reasoning model are explained, and the modeling framework of fuzzy reasoning model is presented. In chapter 4, an empirical study is carried out. Conclusive comments are given at the end of this study.

2. MATHEMATICAL FRAMEWORK OF LATENT CLASS CLUSTERING

A wide variety of applications of latent class clustering are given in McLachlan and Basford (1998) and in McLachlan (1997). In the study, the latent class clustering is applied to classify the heterogeneous data into homogeneous latent classes, and EM algorithm that is widely used algorithms in statistics (Dempster et al., 1977) is presented. The primary advantages of this algorithm are numerical stability, simplicity, and a factorization of the likelihood function (McLachlan and Krishnan, 1997).

It is assumed that the heterogeneous data in our case study has finite mixture distributions, namely, finite homogeneous latent classes. The latent classes are normally distributed with different means and variances. Suppose that the probability density function of a random vector W has a finite mixture of k latent class distribution f_k. The k latent class mixture model has the form

$$f(w;\phi)=\sum_{k=1}^{K}\pi_k\prod_{l=1}^{L}f_{kl}(w_{il};\theta_{kl}) \qquad \text{subject to} \quad \pi_k=1-\sum_{k=1}^{K-1}\pi_k \tag{1}$$

Where $\phi=(\pi_1,\ldots,\pi_{k-1},\theta_{kl})$ is the vector containing the unknown parameters, namely the $k-1$ mixing proportions $\pi_1,\ldots,\pi_{k-1}$. θ_{kl} consists of the parameters of the distribution f_{kl}. L denotes the total number of indicators.

This paper supposes that the k latent class distributions come from multivariate normal densities with unknown means $\mu_{1l},\ldots,\mu_{Kl}$ and unknown variances $\sigma^2_{1l},\ldots,\sigma^2_{Kl}$. Therefore, $\theta_l=(\mu_{1l},\ldots,\mu_{Kl},\sigma^2_{1l},\ldots,\sigma^2_{Kl})$ and

$$f_{kl}(w_{il};\theta_{kl})=\frac{1}{\sqrt{2\pi\sigma^2_{kl}}}\exp\left\{-\frac{(w_{il}-\mu_{kl})^2}{2\sigma^2_{kl}}\right\} \tag{2}$$

The log likelihood function for ϕ that can be formed from the observed data W is given by

$$\log L(\phi)=\sum_{i=1}^{I}\log\left\{\sum_{k=1}^{K}\pi_k\prod_{l=1}^{L}f_{kl}(w_{il};\theta_{kl})\right\} \tag{3}$$

However, the likelihood equation does not yield an explicit solution for ϕ. Therefore, now observed data W is considered to be an incomplete data of a complete data vector $z_i=(z_1,\ldots,z_I)$, where z_i is a k dimensional vector of zero-one indicator variables and where $z_{ik}=(z_i)_k$ is one or zero according as whether w_{il} arose or did not arise from the k^{th} latent class $(k=1,\ldots,K; i=1,\ldots,I)$. If these z_{ik} were observable, then the maximum likelihood estimate of π_k is simply given in Equation (6), which is the proportion of the sample having arisen from the k^{th} latent class. The complete-data log likelihood for ϕ has the multinomial form

$$\begin{aligned}\log L_c(\phi) &= \log\left(\prod_{i=1}^{I}\prod_{k=1}^{K}\left[\pi_k^{z_{ik}}\left\{\prod_{l=1}^{L} f_{kl}(w_{il};\theta_{kl})\right\}^{z_{ik}}\right]\right) \\ &= \sum_{i=1}^{I}\sum_{k=1}^{K}\left\{z_{ik}\log\pi_k + z_{ik}\sum_{l=1}^{L}\log f_{kl}(w_{il};\theta_{kl})\right\} \\ &= \sum_{i=1}^{I}\sum_{k=1}^{K} z_{ik}\log\pi_k + \sum_{k=1}^{K} l_k(\theta_k) \\ \text{subject to} \quad & l_k(\theta_k) = \sum_{i=1}^{I}\left\{z_{ik}\sum_{l=1}^{L}\log f_{kl}(w_{il};\theta_{kl})\right\}\end{aligned} \tag{4}$$

Maximizing the complete-data log likelihood $L_c(\phi)$ is equivalent to maximizing $l_k(\theta_k)$ separately in the partition. Therefore, the EM algorithm requires the iterative Expectation (E) and Maximization (M) of the complete log likelihood $L_c(\phi)$. The M step involves the maximization of a likelihood function that is redefined in each iteration step by the E step. As Equation 4 is linear in the unobservable data z_{ik}, the E-step (on the $(t+1)^{th}$ iteration) simply requires the calculation of the current conditional expectation of z_{ik} given the observed data W.

<E-STEP>
$$E_{\phi'}(z_{ik} \backslash W) = z_{ik}^{(t)} = \pi_k^{(t)}\prod_{l=1}^{L} f_{kl}(w_{il};\theta_{kl}^{(t)}) / \sum_{k=1}^{K}\pi_k^{(t)}\prod_{l=1}^{L} f_{kl}(w_{il};\theta_{kl}^{(t)}) \tag{5}$$

The M-step on the $(t+1)^{th}$ iteration simply requires replacing each z_{ik} by $z_{ik}^{(t)}$ in the equation as follows:

<M-STEP>
$$\pi_k^{(t+1)} = \sum_{i=1}^{I} z_{ik}^{(t)} / I \tag{6}$$

$$\mu_k^{(t+1)} = \sum_{i=1}^{I} z_{ik}^{(t)} w_{il} / \sum_{i=1}^{I} z_{ik}^{(t)} \tag{7}$$

$$\sigma_k^{(t+1)} = \sum_{i=1}^{I} z_{ik}^{(t)} (w_{il} - \mu_k^{(t+1)})^2 / \sum_{i=1}^{I} z_{ik}^{(t)} \tag{8}$$

The EM algorithm starts from an initial solution $\phi^{(t=0)}$ and it develops solution $\phi^{(t=1)},\ldots,\phi^{(t=T)}$ iteratively, where t is the number of iteration $(t = 1,\ldots,T)$. In each iteration, the likelihood value increase 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 Bayesian Information Criterion will be employed to decide the number of latent classes.

$$BIC = -2\ln(L) + p\ln(I) \tag{9}$$

Where p and I are a number of parameters and samples, respectively.

3. ROUTE CHOICE BEHAVIOR MODELS

3.1 Characteristics of Fuzzy Reasoning Model and Random Utility Model

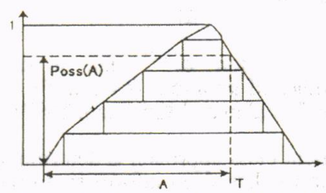
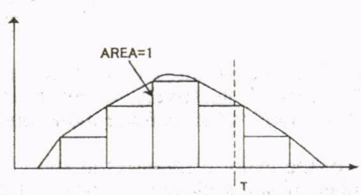
Random utility models have been widely applied to analyze the driver's route choice behavior, and the characteristics and methodologies of the models are well documented. In the other hand, fuzzy reasoning models may be regarded as a new approach in travel choice behavior modeling, even though many related papers have found in recent years. Therefore, the characteristics of fuzzy reasoning models and random utility models are briefly described by comparing the two models.

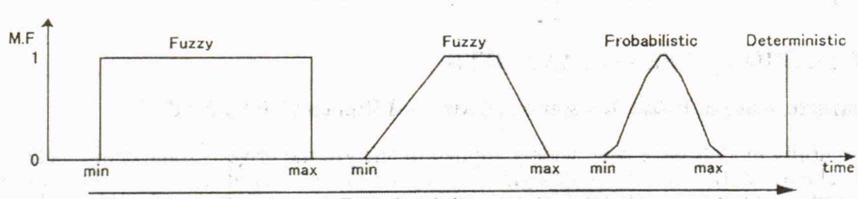
As showed in Table 1, there exist some different characteristics between random utility models and fuzzy reasoning models. In the case of random utility models the non-deterministic nature of driver is mainly focused on the modeling, while the vagueness of driver's perception is focused on the fuzzy reasoning models. In addition, the possibility distribution in the fuzzy reasoning models is applied to consider the knowledge and experience of drivers and experts. For instance, the probability that the travel time less than T is the sum of the probabilities that the time is less than T and each probabilities are independent (see Figure (b) in the table 1). In the other hand, the horizontal strips of Figure (a) in the table 1 are the estimated travel time results by observers, and the results are stacked up. Therefore, the optimal possibility measure is the form

$$Poss(A) = MaxP_{traveltime}(x), \qquad \text{for all } \; x \subseteq A \tag{10}$$

Where A is a set representing "travel time less than T", and $P_{traveltime}(x)$ is possibility distribution (Kikuchi, 1998).

Table 1. Characteristics of fuzzy reasoning models and random utility models

	Fuzzy Reasoning Models	Random Utility Models
Perceptions	Inherent vagueness of driver perceptions	Non-deterministic nature of the route choice behavioral process
Rules	Partially compensatory and partially lexicographic rules of the appraisal of network attributes	Compensatory and lexicographic rules of the appraisal of network attributes
Set	Fuzzy set: [0,1]	Discrete set (Crisp set): {0,1}
Example	A travel time of the route is approximately 20 minutes. The route is now light congestion.	A travel time of the route equals 50 minutes with probability 0.9, while with probability 0.1 the travel time equals 20 minutes
Distribution	 Figure (a) Possibility distribution	 Figure (b) Probability distribution
Estimation	Trial-and-error estimation	Maximum likelihood estimation
Criteria	Goodness of Fit	ρ^2, Goodness of Fit



Figure(c) Change of driver's perception on the network attributes

(Hoogendoorn et al, 1999; Kikuchi, 1998; Lotan and Koutsopoulos, 1993[a,b])

However, some drawbacks of the fuzzy reasoning models are criticized in behavior modeling. One is the estimation problem of the parameters describing the membership functions of fuzzy reasoning models and of the defuzzification yielding a representative value of the fuzzy reasoning results. There are no exact methodological strategies yet. Another is a criterion of the fuzzy reasoning models. The goodness of fit has been used to assess a quality of the estimated model, but a higher goodness of fit does not means a better model (Hoogendoorn et al, 1999).

The Figure (c) in the bottom line of Table 1 instances the changes in the driver's perceptions on the road networks as he or she is more familiar with the road networks by gaining wide experience; the membership functions are becoming a more narrow range and less diffuse shape. Their route choice behaviors correspond on the probability distribution (Lotan and Koutshopoulos, 1993).

3.2 Modeling Framework of Fuzzy Reasoning Model

This chapter outlines the modeling framework of fuzzy reasoning models in detail, while the explanation on the random utility models is abbreviated to the simplicity of the paper.

Fuzzy IF-THEN rules, in which the antecedents and the consequent involve linguistic variables, are applied to model the decision-making process of drivers. The general rule form of this paper is

$$R^k : \text{If x is } A_k, \text{AND, y is } B_k, \text{THEN } z = C_k, \qquad k = 1,...,K \qquad (11)$$

Where x, y and z are linguistic variables representing the input variables and the control variable respectively. A_k, B_k and C_k are the linguistic predicates of the linguistic variables x, y and z in the universes of discourse U, V and W respectively. k is the number of fuzzy rules. In this paper, A_k and B_k characterize the ordinary travel time levels of each alternative route.

Travel time = {short, moderate, long} = {short travel time, moderate travel itme, long travel time}.

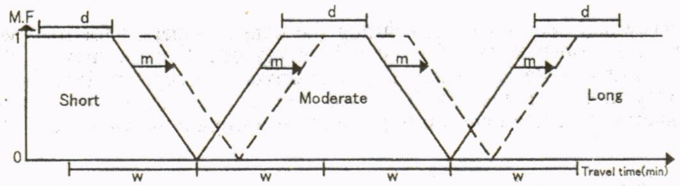


Figure 1. Membership functions of travel time

In this paper, the optimal fuzzy membership functions are estimated by trial-and-error method. Three indices shown in Figure 1 are employed for the estimation. Index "m" is to approach the membership functions into the real perceptions of drivers on the travel times, "w" is to find the optimal overlap range of membership functions, and "d" is to estimate the level of vagueness of drivers. The goodness of fit are estimated at the each step in which the three indices are changed, and finally the optimal membership functions are decided at the point of the highest goodness of fit. Note that the initial points of membership functions can be assumed at any point of collected data, and the overlap rate of membership functions begin from 0%.

For the consequent term, five linguistic predicates characterize the preference levels of drivers on the alternative routes as shown in Figure 2: C_k (preference)= {HPH, PH, M, PL, HPL} where the predicates mean High Preference of Highway, Preference of Highway, Middle Preference, Preference of Local way, and High Preference of Local way, respectively.

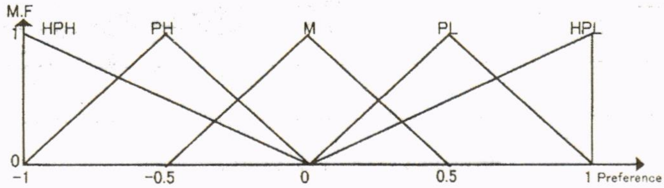


Figure 2. Membership functions of preference

Finally, nine fuzzy inference rules are established as shown in Table 2, and all rules will be fired in parallel. For estimating fuzzy reasoning models, Min-Max composition strategies is employed, and Centroid of Area method is applied as a defuzzification method to extract a crisp value that represents the possibility distribution of an inferred fuzzy linguistic value set (Jang et al, 1997).

Table 2. Fuzzy inference rules

		Perceived highway travel time		
		Short	Moderate	Long
Perceived local way travel time	Short	M	PL	HPL
	Moderate	PH	M	PL
	Long	HPH	PH	M

4. EMPIRICAL STDUY

4.1 The Data

In 1996, the survey was conducted at two intercity roads, Honam highway and No.22 local way among driving commuters from Sunchon city to Kwangju metropolitan with an objective to examine perceived levels of travel time on the alternative routes and ordinary choice route. The number of 504 sheets was totally distributed at the both roadsides and collected by examiner. The available response rate was 56.35%(284/504). To survey the perceived levels of travel time, the questionnaire was designed to ask three perceived levels of travel time of each alternative route. For instance, "How long travel time does the route takes, if the route takes a short time, a moderate time, and a long time?"

Note that the survey had been conducted during the construct period for road widening in Honam highway. The construction of Honam highway could cause making the driver's perception on travel time more vague travel time of driver's perception since the travel time of Honam highway might be unstably increased due to the construction. Therefore, this situation of the alternative route gives an opportunity to estimate the vagueness of drivers in our case study.

The ordinary choice rates of each alternative routes was 45%(129/284) and 55%(155/284) at highway and local way, respectively. The reason is regarded that the respondents couldn't get enough compensation for their toll due to the construction of Honam highway. The variations of ordinary choice rates of highway are displayed in Figure 3. As the difference of perceived levels of travel time is bigger, the choice pattern is more distinct. The result supports that route choice behaviors of drivers in this case study are mainly influenced by their perceived levels of travel time.

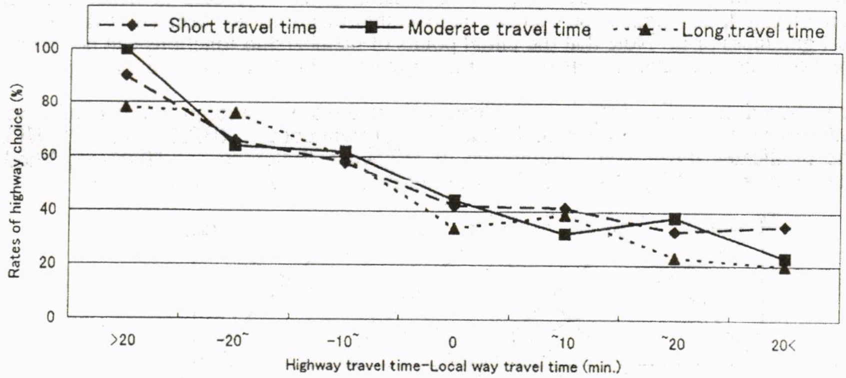


Figure 3. Variations of choice rate considering the perceived levels of travel time

4.2 Latent Class Clustering Estimation

Latent class clustering approach is applied to classify a heterogeneous population of sample data into finite homogeneous subpopulations, and travel choice behavior models are estimated by using each homogeneous subpopulation. For latent class clustering, it is assumed that if the difference of perceived levels of travel time is big, the driver's choice behavior is based on randomness uncertainty so that random utility models are appropriate to analyze his/her route choice behavior. In the other hand, if the difference of perceived levels is small, the driver's choice behavior is based on vagueness uncertainty. Therefore, fuzzy reasoning models are available.

Table 3 shows BIC values for the estimated one to four latent class clustering, and one latent class model corresponds to no heterogeneity of the data. The 2-latent class model has the lowest BIC value, which indicates slightly lower than the 3-latent class model. Therefore, the classification results of 2-latent classes are employed in this study.

Table 3. Latent class clustering

Num. Of Class	Num. OF parameters	BIC
1-Latent class	6	3081.9
2-Latent classes	13	2614.7
3-Latent classes	20	2808.2
4-Latent classes	27	7732.9

The estimated parameters of the 2-latent classes model are reported in Table 4. The latent class 1 has much higher means and variances than those of the latent class 2, and the result represents that the drivers of latent class 1 have more distinct perceived levels of travel time on the alternative routes than the drivers of latent class 2. Therefore, it is assumed that the route choice behaviors of drivers in latent class 1 are based on randomness uncertainty, while the route choice behaviors of drivers in latent class 2 are based on vagueness uncertainty.

Table 4. 2-latent classes clustering

Latent class	Latent class 1 (49.2%)		Latent class 2 (50.8%)	
Variables	μ_1	σ_1	μ_2	σ_2
Short travel time	3.143	1.512	1.595	0.525
Moderate travel time	3.179	1.451	1.449	0.523
Long travel time	3.930	1.629	1.569	0.527

4.3 Route Choice Behavior Models

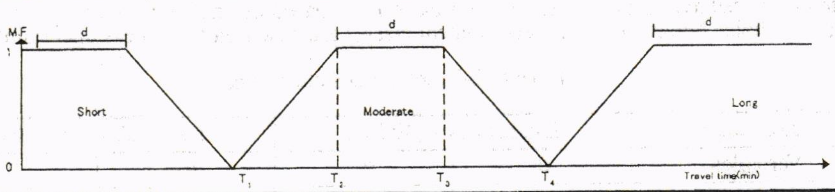
Table 5 contains the estimated results of random utility models based on each latent class and whole data. For the travel time variable of random utility models, the average value of three perceived levels of travel time is employed due to the multicollinearity of the model. Age (if more than 34 years old is 1, otherwise 0) and wiliingness to switch by traffic situations (if switch is 1, otherwise 0) are added for the logit model. The highest estimation result in random utility models is outputted in the latent class 1 (randomness class), while the lowest result is in the latent class 2 (vagueness class). Moreover, the coefficient value of willingness to switch variable shows that respondents of the randomness class are more sensitive to traffic situations than respondents of the vagueness class. The opposite estimation results are in the fuzzy reasoning models, that is, the goodness of fit of latent class 2 is much higher than that of latent class 1 and little higher than that of whole data. Moreover, the goodness of fit of latent class 2 is highly improved by employing the fuzzy reasoning model (from 62.76% to 63.45%). The "d" index in Table 6 shows the level of vagueness uncertainty, and the most vagueness class is the latent class 2. These results represent that the vagueness of uncertainty is well considered in the fuzzy reasoning model than the random utility model. In the other hand, the random utility model is more applicable to treat the randomness of uncertainty than the fuzzy reasoning model. Eventually, all results in table 5 and 6 support that the two modeling methodologies can dealt the uncertainty of travel time with a certain amount of reliability. However, the accuracy of route choice behavior model can be reduced, if the heterogeneity of sample data is not considered for modeling In addition the necessity of a combined model of random utility model and fuzzy reasoning model is emphasized from the results. Concerning on the combined model, the latent class clustering (or latent class model) may be the core of the methodological strategies since the expected probability to include the each latent class can be obtained in the analysis, Equation (5).

Table 5. Random utility models

	Whole data	Latent class 1	Latent class 2
Variables	Coefficient (t-value)		
Travel time	-0.044 (-4.32) **	-0.045 (-3.95) **	-0.055 (-2.01) *
Age (>35)	-0.024(-2.59) **	-0.021(-1.51)	-0.025(-1.97) *
Switch	0.640(2.48) *	0.714(1.77)	0.570(1.62)
# Of sample	284	139	145
$L(0)-L(\beta)$	-16.414	-11.816	-5.599
ρ^2	0.083	0.123	0.056
Adjusted ρ^2	0.068	0.092	0.026
Goodness of fit	64.79%	66.91%	62.76%

Note)**: Significant at the 0.01 confidence level, *: Significant at the 0.05 confidence level

Table 6. Fuzzy reasoning models



Variables		T_1	T_2	T_3	T_4	d	Goodness of fit
Whole Data							
Highway	Short	0	30	42	49	12	62.676%
	Moderate	45	54	66	75		
	Long	71	79	150	150		
Local way	Short	0	30	60	78		
	Moderate	72	84	96	108		
	Long	102	120	150	150		
Latent Class 1							
Highway	Short	0	30	30	48	2	61.871%
	Moderate	41	59	61	79		
	Long	72	90	150	150		
Local way	Short	0	30	57	72		
	Moderate	69	84	86	101		
	Long	98	113	150	150		
Latent Class 2							
Highway	Short	0	30	40	49	20	63.448%
	Moderate	41	50	70	79		
	Long	71	80	150	150		
Local way	Short	0	30	70	79		
	Moderate	71	80	100	109		
	Long	101	110	150	150		

5 CONCLUSIONS

The main purposes of the paper are to consider the heterogeneity of driver perceptions in sample data by using the latent class clustering approach, to estimate random utility models and fuzzy reasoning models considering the heterogeneity, and to find an appropriate model by the types of uncertainty. The results are summarized as follows: Firstly, the heterogeneous data in our case study can be classified into finite homogeneous data by employing the latent class clustering approach. Secondly, the estimation results of random utility models and fuzzy reasoning models show that if the travel choice behavior models are developed without considering the type of uncertainty, the accuracy of models may be deteriorated. Finally, the necessity of a combined model considering the randomness and the vagueness uncertainty simultaneously is emphasized as a future work.

The goodness-of-fit indices of random utility model were not sufficient in this study, because of the limitation of data availability on perceived travel time. A more efficient survey method collecting individual's perceptions on travel attributes must be developed in the following studies.

ACKNOWLEDGEMENT

This research was supported by grand No. 2000-2-31300-001-3 from the basic research program of the Korea Science and Engineering Foundation.

REFERENCES

Adler, J. and Casello, J. (1997) A two-stage fuzzy control methodology for dynamic routing under ATIS, **Workshop on Dynamics and ITS response**, The 8th meeting of IATBR '97, Austin, Texas.

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, 39,** 1-38.

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.

Hoogendoorn, S., Hoogendoorn-Lanser, S. and Schuurman, H. (1999) Fuzzy perspectives in traffic engineering, **Workshop on Intelligent Traffic Management Models**, Delft, Netherlands.

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

Kikuchi, S. (1998) Treatment of uncertainty in study of transportation: fuzzy set theory and evidence theory, **Journal of Transportation Engineering, Vol. 124, No.1**.

Lin, C.T. and George Lee, C.S. (1996) Neural fuzzy systems. Prentice-Hall, Inc.

McLachlan, G.J. and Krishnan, T. (1997) **The EM Algorithm and Extensions.** John Wiley and Sons, Inc.

Lotan, T. and Koutsopoulos, H. (1993[a]) Approximate reasoning models for route choice behavior in the presence of information, **Transportation and Traffic Theory,** 71-88.

Lotan, T. and Koutsopoulos, H. (1993[b]) Models for route choice behavior in the presence of information using concepts from fuzzy set theory and approximate reasoning, **Transportation 20,** 129-155.

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