THE EFFECTIVENESS OF NON-IIA MODELS ON TIME CHOICE BEHAVIOR ANALYSIS

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

Yusuke KANDA
Oriental Consultants Co.,Ltd.,
3-5-7 Hisamoto, Takatsu-ku, Kawasaki
Kanagawa, 213-0011, Japan
Fax: +81-3-3409-7551
E-mail: yusuke77@mx91.tiki.ne.jp

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

Toshiyuki OKAMURA
Research Associate
Graduate School for IDEC
Hiroshima University
Kagamiyama 1-5-1, Higashi-Hiroshima, 739-8529, Japan
Fax: +81-824-24-6922
E-mail: tokamura@hiroshima-u.ac.jp

Abstract: This paper aims to confirm the existence of similarity and heteroscedasticity among alternatives and to examine the effectiveness of new discrete choice models that can relax the IIA property of MNL model with the context of time choice. As an empirical study, two non-IIA models are estimated to describe the arrival time choice behavior under the flex-time working hours, those are Paired Combinatorial Logit (PCL) model and Heteroscedastic Extreme Value (HEV) model. The continuous arrival time is categorized into three different discrete alternatives with various time periods and time lengths of alternatives in order to observe the influence of time categorizes. The estimation results of the PCL and HEV models show that the similarity and heteroscedasticity are ineligible in case of time choice.

Key Words: Travel demand forecasting, IIA Property, PCL model, HEV model

1. INTRODUCTION

MNL (Multinomial Logit) model, based on the random utility maximization theory, has a very simple and intuitive form of choice probability. It leads to the fact that the MNL model is often applied to travel behavior analysis and travel demand forecasting, because of its computational convenience. The MNL model supports the restrictive IIA (Independence of Irreverent from Alternatives) property based on some assumptions.

One of the assumptions is the IID assumption, which assumes the random components of utility functions of different alternatives are independent and identically distributed with Gumbel. Hence, it has been pointed out that the biases in estimated parameters arise frequently.

Consider the modeling of travel time choice behavior. The curve fitting methods including regression models and duration models have often been employed to describe the observed distribution, while the individual decision making process cannot be tackled with any theories such as random utility maximization, and transport policies are insensitively evaluated based on the models. This is why discrete choice models are commonly used to predict the time choice behavior.

The sequential time categories seems dependent each other, because the levels of services are variant in series and are almost equivalent between the moments before and after a boundary of time categories. Moreover, the differences between estimated and observed time choice results may strongly depend on the definition of time categories. If most respondents face the above boundary problem, the errors will not be identical and independent across the alternatives.

In this paper, we aim to confirm the existence of similarity and heteroscedasticity among time choice alternatives and to examine the effectiveness of new discrete choice models that can relax the IID assumption of MNL model on time choice. First, in Chapter 2, we review and summarize non-IIA models recently developed. Secondly, we explain two non-IIA models, namely PCL model and HEV model in detail. Finally, in chapter 4, we apply the two models to analyze time choice behavior to examine their effectiveness.

2. DISCRETE CHOICE MODELS WITH FLEXIBLE ERROR STRUCTURE

Various models have recently been developed in order to relax the IID assumption fully or partially. These models are roughly classified into three categories by the following relaxing approach; 1) relaxing the assumption of independent and identical distribution, 2) relaxing the assumption of independent distribution, and 3) relaxing the assumption of identical distribution.

(1) Models relaxing the assumption of independent and identical distribution

The most general models, without any restrictive assumption on error structure, are the MNP (Multinomial Probit) model and the MXL (Mixed Logit) model.

The MNP model allows a flexible structure for the covariance among the random components of the alternatives. However, this model requires high-dimensional multivariate normal integration of the order of the number of the alternatives in the choice probability expressions. Several efficient simulation methods, which approximate the high-dimension integration, have been developed since 90's. (eg. Yai et.al. 1997)

The MXL model is a more flexible logit model with random-coefficients, but does not support the IIA property. The various distributions of the coefficients in utility function can provide not only heterogeneity over respondents, but also correlation and heteroscedasticity among alternatives.

Mass Point Logit (MPL) model which was proposed by Sugie et al. (1999) is also a kind of MXL model, in which several mass points separate the error distribution.

(2) Models relaxing the assumption of independent distribution

All models in this group belong to GEV model family (McFadden, 1978). The advantage of the GEV models is to maintain closed-form expressions for choice probabilities.

The simplest model that permits covariance in error components is called as NL (Nested Logit) model. The NL model has a logsum parameter that determines the correlation in unobserved components among alternatives in a nest, while the model cannot deal with the correlations among alternatives in the other nests.

PCL (Paired Combinatorial Logit) model has a more general form of correlations among alternatives than NL model. This model allows differential correlations across all pairs of alternatives. The degrees of correlations are measured by unknown similarity parameters.

CNL (Cross-Nested Logit) model has also more general. In this model, an alternative needs not be exclusively assigned to one nest as in the nested logit structure, and an alternative appears in multiple nest in a same level with different weights defined as allocation parameters.

GNL (Generalized Nested Logit) model has the most general structure, because of combining the above characteristics of the PCL and CNL models. The model has similarity parameters among possible all pairs of alternatives and allocation parameters of all nests. The GNL model corresponds to PCL model when all allocation parameters are constrained to be equal, while the model also becomes CNL model when all similarity parameters are constrained to be equal.

The NL, CNL, and GNL models all require a priori specification of the hierarchical nest structure. More alternatives respondents have, more possible structures exist. However, only PCL model can avoid such a burdensome task. The PCL is the most suitable since it is difficult for analysts to predetermine the appropriate correlation structure. Consequently, the PCL model is employed in the rest of this study.

Table 1. Models with non-independent and non-identical error distribution

Multinomial Probit (MNP) model (Daganzo, 1979)

$$P_i = \int_{\eta_{i1}=-\infty}^{V_i - V_1} \cdots \int_{\eta_{i1}=-\infty}^{V_i - V_2} \cdots \int_{\eta_{im}=-\infty}^{V_i - V_m} \cdots \phi(\boldsymbol{\eta}) d\eta_{im} \cdots d\eta_{i2} d\eta_{i1} \qquad (1)$$

$$\phi(\boldsymbol{\eta}) = (2\pi)^{-(m-1)/2} |\mathbf{S}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\eta}'\mathbf{S}^{-1}\boldsymbol{\eta}\right) \qquad (2)$$

$$\eta_{ij} = \varepsilon_j - \varepsilon_i \qquad (3)$$

P_i : choice probability of alternative i
V_i : observed utility of alternative i
ε_i : random utility of alternative i
$\mathbf{S}$: error variance-covariance matrix

Mixed Logit (ML) model (McFadden et. al., 1997)

$$P_i = \int L_i(\eta) f(\eta \mid \Omega) d\eta \qquad (4)$$

$$L_i = \frac{\exp(V_i + \eta_i)}{\sum_j \exp(V_j + \eta_j)} \qquad (5)$$

η: random realization parameter
Ω: vector of underlying moment parameters
f : probability density function

Mass Point Logit (MPL) model (Sugie et. al., 1999)

$$P_i = \sum_{k=1}^{m} \frac{\exp(V_i + \xi ki)}{\sum_{i'=1}^{I} \exp(Vi' + \xi ki')} \rho_k \qquad (6)$$

$$s.t. \quad \sum_{k=1}^{m} \rho_k = 1 \quad \rho_k \geq 0$$

ξ : location parameter
ρ : weight parameter
k number of mass points

Table 2. Models with non-independent but identical error distribution

Nested Logit (NL) model (McFadden, 1978)

$$P_{ijk} = \frac{\exp\left(\dfrac{V_{ijk}}{\varpi_{jk}}\right)}{\sum_{i'}\exp\left(\dfrac{V_{i'jk}}{\varpi_{jk}}\right)} \times \frac{\exp\left(\dfrac{\varpi_{jk}}{\varpi_k}\times\Gamma_{jk}\right)}{\sum_{j'}\exp\left(\dfrac{\varpi_{j'k}}{\varpi_k}\times\Gamma_{j'k}\right)} \times \frac{\exp(\varpi_k\times\Gamma_k)}{\sum_{k'}\exp(\varpi_{k'}\times\Gamma_{k'})} \tag{7}$$

$$\Gamma_{jk} = \ln\sum_{i'}\exp\left(\frac{V_{i'jk}}{\varpi_{jk}}\right); \text{ Logsum parameter} \tag{8}$$

P_{ijk} : Choice probability of alternative i in nest j and k (j; lower nest)

ϖ : Structure parameter (Similarity parameter)

Paired Combinatorial Logit (PCL) model (Chu, 1989)

$$P_i = \sum_{j\neq i}\left[\frac{(\alpha\exp(V_i))^{\frac{1}{\varpi_{ij}}}}{(\alpha\exp(V_i))^{\frac{1}{\varpi_{ij}}} + (\alpha\exp(V_j))^{\frac{1}{\varpi_{ij}}}} \times \frac{\left((\alpha\exp(V_i))^{\frac{1}{\varpi_{ij}}} + (\alpha\exp(V_j))^{\frac{1}{\varpi_{ij}}}\right)^{\varpi_{ij}}}{\sum_{k=1}^{J-1}\sum_{m=k+1}^{J}\left((\alpha\exp(V_k))^{\frac{1}{\varpi_{km}}} + (\alpha\exp(V_m))^{\frac{1}{\varpi_{km}}}\right)^{\varpi_{km}}}\right] \tag{9}$$

J: Number of alternatives

α : Allocation parameter

Cross-Nested Logit (CNL) model (Vovsha, 1997)

$$P_i = \sum_m P_{i|m}\times P_m = \sum_m\left[\frac{(\alpha_{im}\exp(V_i))^{\frac{1}{\varpi}}}{\sum_{j\in N_m}\left(\alpha_{jm}\exp(V_j)\right)^{V_i\frac{1}{\varpi}}} \times \frac{\left(\sum_{j\in N_m}(\alpha_{jm}\exp(V_j))^{\frac{1}{\varpi}}\right)^{\varpi}}{\sum_m\left(\sum_{j\in N_m}(\alpha_{jm}\exp(V_j))^{\frac{1}{\varpi}}\right)^{\varpi}}\right] \tag{10}$$

$$s.t \quad \sum_m \alpha_{im} = 1, \quad \forall i\alpha_{im} > 0, \quad \forall i, m$$

N_m : Set of all alternatives included in nest m

Generalized Nested Logit (GNL) model (Wen et. al, 2000)

$$P_i = \sum_m P_{i|m}\times P_m = \sum_m\left[\frac{(\alpha_{im}\exp(V_i))^{\frac{1}{\varpi_m}}}{\sum_{j\in N_m}(\alpha_{jm}\exp(V_j))^{\frac{1}{\varpi_m}}} \times \frac{\left(\sum_{j\in N_m}(\alpha_{jm}\exp(V_j))^{\frac{1}{\varpi_m}}\right)^{\varpi_m}}{\sum_{m'}\left(\sum_{j\in N_{m'}}(\alpha_{jm'}\exp(V_j))^{\frac{1}{\varpi_{m'}}}\right)^{\varpi_{m'}}}\right] \tag{11}$$

$$s.t \quad \sum_m \alpha_{im} = 1, \quad \forall i\alpha_{im} > 0, \quad \forall i, m$$

Table 3. Models with independent but non-identical error distribution

Heteroscedastic Extreme Value (HEV) model (Bhat, 1995)
$$P_i = \int_{\varepsilon=-\infty}^{\varepsilon=+\infty} \prod_{j \neq i, j \in C} F\left[\theta_j (V_i - V_j + \varepsilon_i)\right] \theta_i f(\theta_i \varepsilon_i) d\varepsilon_i \quad (12)$$ F cumulative density function θ_i scale parameter of alternative I
Oddball Alternative model (Recker, 1995)
$$P_i = \frac{\exp(V_i + \eta_i)}{\sum_{\forall \ell = A, t \neq i} \exp(V_\ell + \eta_\ell)} \left[1 - \phi_r \exp(\phi_r)^{\phi r} EI(\phi_r)\right] \quad \forall i \in J, i \neq r, \quad (13)$$ $$\phi_r = \frac{\exp(V_i + \tilde{V}_r + \eta_i + \tilde{\eta}_r)}{\sum_{\forall \ell = J, t \neq k} \exp(V_\ell + \eta_\ell)} \quad (14)$$

(3) Models relaxing the assumption of identical distribution

Several models that allow non-identical random components have been proposed. HEV (Heteroscedastic Extreme Value) model is an alternative of these discrete choice models. The main virtues of the HEV model are its allowance of different variance of error term across alternatives. Unlike the NL model, the model does not require the prior partitioning of the choice set into mutually exclusive branches

Oddball Alternative model permits the random utility variance of one "oddball" alternative to be larger than those of other alternatives. This model has a closed-form for choice probabilities. However, it is restrictive in requiring that all alternatives except the one oddball have identical variance. Therefore, the HEV model using in this study is more flexible rather than the Oddball Alternative model.

3. MODEL CHARACTERISTICS

a) PCL model

PCL model is derived from GEV function by referring Chu (1989). The GEV model can be derived from following function.

$$G(Y_1, Y_2, \cdots, Y_n), \quad Y_1, Y_2, \cdots, Y_n \geq 0 \quad (15)$$

which is non-negative, homogeneous of degree one, approached infinity with any Y_i, $i = 1, 2, \ldots, n$ and has kth cross-partial derivatives which are non-negative for odd k and non-positive for even k. Suppose a given function which satisfies such three conditions defines a probability function for alternative i as

$$P_i = \frac{Y_i G_i(Y_1, Y_2, \cdots, Y_n)}{G(Y_1, Y_2, \cdots, Y_n)}, \quad (16)$$

where G_i is the first derivatives of G with respect to Y_i. The transformation, $Y_i = \exp(V_i)$, where V_i represents the observable components of the utility for alternative i, is used to ensure positive Y_i. The PCL model is derived from the function G:

$$G(Y_1, Y_2, \cdots, Y_n) = \sum_{i=1}^{n-1} \sum_{i \neq j}^{n} \left\{ Y_i^{1/1-\sigma_{ij}} + Y_j^{1/1-\sigma_{ij}} \right\}^{1-\sigma_{ij}} \tag{17}$$

where the double summation includes all pairs of alternatives in the choice set and σ_{ij} is an index of the similarity between alternative i and j. The PCL model is consistent with random utility maximization if the conditions, $0 \leq \sigma < 1$, are satisfied for all pairs. If $\sigma_{ij} = 0$ for all alternative pairs, the PCL model collapses to the MNL model.

Substituting equation Eq. (17) into Eq. (16) and using the transformation of Y_i gives the probability of choosing alternatives i as

$$P_{ij} = \frac{\sum_{j \neq i} \exp\left(\frac{V_i}{1-\sigma_{ij}}\right) \left\{ \exp\left(\frac{V_i}{1-\sigma_{ij}}\right) + \exp\left(\frac{V_j}{1-\sigma_{ij}}\right) \right\}^{-\sigma_{ij}}}{\sum_{q=1}^{n-1} \sum_{q=r+1}^{n} \left\{ \exp\left(\frac{V_q}{1-\sigma_{qr}}\right) + \exp\left(\frac{V_r}{1-\sigma_{qr}}\right) \right\}^{1-\sigma_{qr}}} \tag{18}$$

Eq. (18) corresponds to Eq. (9) by substituting ϖ_{ij} into $1-\sigma_{ij}$ and by canceling out α. This expression can be rewritten as

$$P_i = \sum_{j \neq i} P_{i|ij} P_{ij} \tag{19}$$

$$= \sum_{j \neq i} \frac{\exp\left(\frac{V_i}{1-\sigma_{ij}}\right)}{\exp\left(\frac{V_i}{1-\sigma_{ij}}\right) + \exp\left(\frac{V_j}{1-\sigma_{ij}}\right)} \times \frac{\left\{ \exp\left(\frac{V_i}{1-\sigma_{ij}}\right) + \exp\left(\frac{V_j}{1-\sigma_{ij}}\right) \right\}^{1-\sigma_{ij}}}{\sum_{q=1}^{n-1} \sum_{q=r+1}^{n} \left\{ \exp\left(\frac{V_q}{1-\sigma_{qr}}\right) + \exp\left(\frac{V_r}{1-\sigma_{qr}}\right) \right\}^{1-\sigma_{qr}}}, \tag{20}$$

where $P_{i|ij}$ is the conditional probability of choosing alternative i given the chosen binary pair and P_{ij} is the unobserved probability for the pair of alternative i and j.

b) HEV model

HEV model supposes that the random component of utility function is distributed independently but not identically. The CDF for each random component is the type I extreme distribution with zero mean and scale parameter θ_i. Hence, the CDF of the random error term and the choice probability of alternative i are written as

$$F(\varepsilon) = \exp(-\exp(-\theta_i \varepsilon_i)), \tag{21}$$

$$P_i = \int_{\varepsilon=-\infty}^{\varepsilon=+\infty} \prod_{j \neq i, j \in C} F\left[\theta_j (V_i - V_j + \varepsilon_i)\right] \theta_i f(\theta_i \varepsilon_i) d\varepsilon_i \tag{22}$$

The probabilities and derivatives must be evaluated numerically, as there is no closed form for the integral. This model has one-dimensional integral, so that Gauss-Laguerre quadrature is required for parameter estimation. (Bhat, 1995)

4. EMPIRICAL STUDY

(1) Data Characteristics

A questionnaire survey was conducted for full time workers in the office, which has introduced the flex-time working hours in Hiroshima, 1997, in order to analyze workers' commuting behavior under the working hours. The number of effective responses came to 301; that was 96.2% of whole samples.

Figure 1 shows the observed distribution of workers' arrival time at office. The respondents fulfilled the perceived arrival time in minute. The average of office arrival time indicates 8:56 a.m., which is in the peak of the distribution.

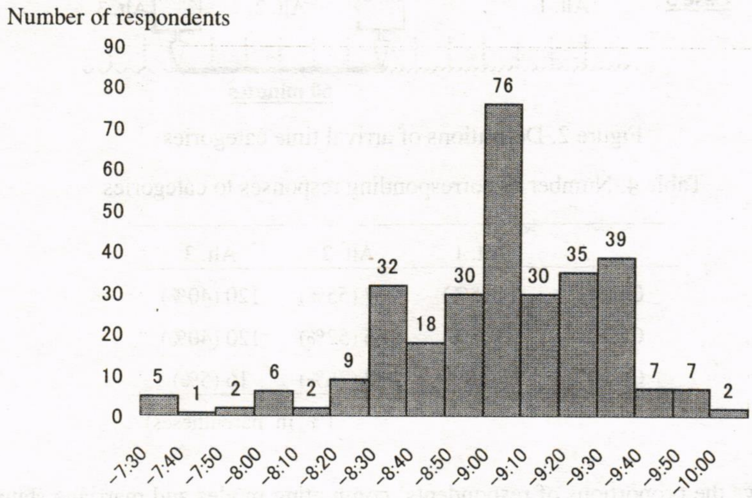

Figure 1. Distribution of observed arrival time

For the sake of applying discrete choice models to this arrival time choice data, the following subordinate works are needed. First, the reported continuous arrival time is sliced off into three discrete categories. Three different cases with the combination of various time periods and time lengths of alternatives are defined in order to investigate the influence of time categorizes on the similarity and heteroscedasticity (Fig. 2).

The alternatives 2 and 3 have a common boundary of categories at 9:00, which is equal to the peak of the responded time distribution in Cases 1 and 2, while they have a different boundary at 9:30 in Case 3. Stronger similarity might appear between the error components of the alternatives 2 and 3 in Case 1 and 2, since many responses would shift to another sequential alternative if the time is re-categorized only a few minutes later. Moreover, to compare the influence of difference of time length, the range of the alternative 2 in Case 1 is set up to 60 minutes, whereas 40 minutes in Case 2.

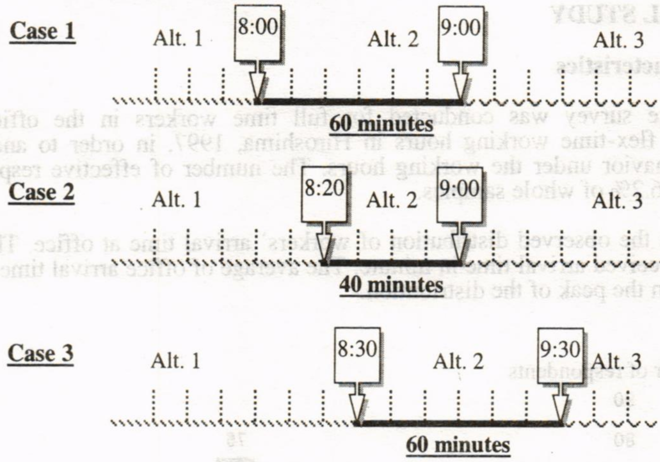

Figure 2. Definitions of arrival time categories

Table 4. Number of corresponding responses to categories

	Alt. 1	Alt. 2	Alt. 3
Case1	14 (5%)	167 (55%)	120 (40%)
Case2	25 (8%)	156 (52%)	120 (40%)
Case3	57 (19%)	228 (76%)	16 (5%)

(% in parentheses)

Figure 3 shows the proportions of respondents' commuting modes and marriage status. More than half of workers travel to their office by car, and the other 20 percent of workers use public transportation (train, bus or tram). Concerning the individual marriage status, less than 20 percent of workers are single, so that a fifth of workers have no constraint on time choice by their household members.

The relationship between the workers' arrival time and age is shown in Figure 4. Younger workers tend to arrive at the office later.

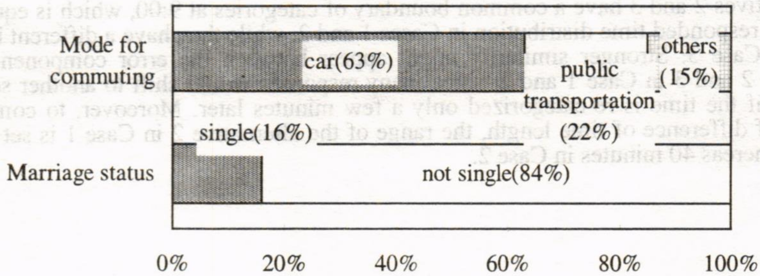

Figure 3. Commuting mode and household

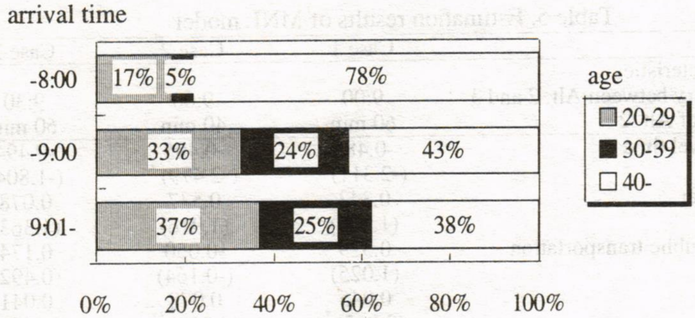

Figure 4. Relationship between office arrival time and age

Figure 5 shows the relationship between office arrival time and commuting travel time. Workers who have shorter travel time tend to arrival at the office later, because they have less risk of travel delay. It is found that individual socio-demographic characteristics and the level of travel services significantly affect on the time choice behavior.

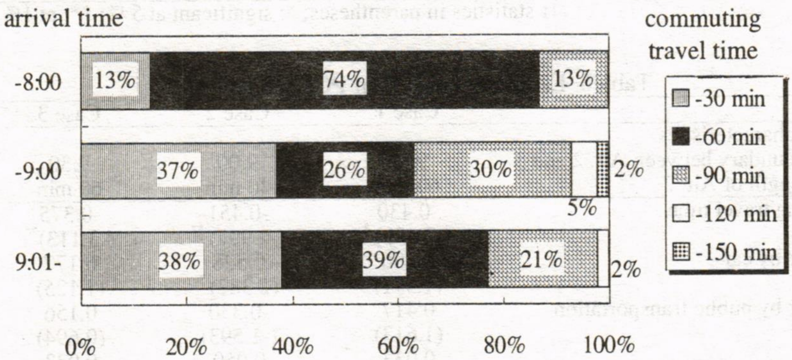

Figure 5. Relationship between office arrival time and commuting travel time

(2) Model Estimation and Discussions

The PCL and HEV models are estimated in three different cases in order to examine the influences of time categories on the similarity and heteroscedasticity of alternatives in Figure 2. The estimation results of the MNL, PCL and HEV models are shown in Table 5, 6 and 7, respectively. These models can be estimated by using MLE, similar to the conventional MNL models. It is not required to employ any complicated simulation procedure.

The PCL and HEV models are superior to the MNL model in terms of the log-likelihood ratio. Some estimated parameters show inconsistent effects (ie. unexpected positive or negative signs of parameters). For instance, the estimated parameters of marriage status 'single' dummy are positive in Cases 1 and 2, but negative in Case 3, and the signs of parameters of 'commuter by public transportation' dummy are variant over cases. These results turn out that some public transportation users and single workers have a tendency to be included in different alternatives by different definitions of time categories.

Regarding on the estimated results of similarity parameters, all similarity parameters excluding σ_{12} in Cases 2 and 3 are statistically significant. It is, therefore, found the time categories are not independent.

Table 5. Estimation results of MNL model

Variable	Case 1	Case 2	Case 3
Category characteristics			
Time boundary between Alt. 2 and 3	9:00	9:00	9:30
Time length of Alt. 2	60 min	40 min	60 min
Commuting travel time	-0.485	-0.495	-0.494
(hour)	(-2.311)*	(-2.479)*	(-1.804)
Commuter by car	0.343	0.537	0.078
(dummy)	(1.200)	(1.916)	(1.363)
Commuter by public transportation	0.329	-0.050	0.174
(dummy)	(1.025)	(-0.164)	(0.492)
Age	0.045	0.078	0.041
(year)	(2.065)*	(4.360)**	(3.120)**
Single	0.044	0.056	-0.453
(dummy)	(0.127)	(0.160)	(-1.094)
Constant for Alt. 2	4.290	5.149	3.050
	(3.522)**	(5.897)**	(5.265)**
Constant for Alt. 3	3.907	4.720	0.297
	(3.184)**	(5.419)**	(0.482)
Initial Log-likelihood	-330.680	-330.680	-330.680
Maximum Log-likelihood	-244.348	-259.324	-197.010
Adjusted Log-likelihood ratio	0.251	0.205	0.396
Number of samples	301	301	301

(t-statistics in parentheses; *: significant at 5 %; **: at 1%)

Table6. Estimation results of PCL model

Variable	Case 1	Case 2	Case 3
Category characteristics			
Time boundary between Alt. 2 and 3	9:00	9:00	9:30
Time length of Alt. 2	60 min	40 min	60 min
Commuting travel time	-0.430	-0.451	-0.375
(hour)	(-2.499)*	(-2.997)**	(-3.113)**
Commuter by car	0.306	0.528	0.177
(dummy)	(1.591)	(2.382)*	(1.125)
Commuter by public transportation	0.417	-0.330	0.156
(dummy)	(1.613)	(-1.593)	(0.694)
Age	0.033	0.050	0.033
(year)	(2.308)*	(3.861)**	(3.448)**
Single	0.061	0.124	-0.403
(dummy)	(0.372)	(0.442)	(-3.104)
Constant for Alt. 2	3.537	3.526	2.520
	(5.517)**	(5.053)**	(6.046)**
Constant for Alt. 3	3.543	3.816	1.183
	(5.802)**	(5.768)**	(2.596)**
Similarity parameter			
σ_{12}	0.126	0.252	0.105
	(2.578)**	(1.715)	(0.673)
σ_{13}	0.414	0.787	0.630
	(3.778)**	(4.600)**	(3.058)**
σ_{23}	0.909	0.855	0.690
	(13.427)**	(12.888)**	(2.518)**
Initial Log-likelihood	-330.680	-330.680	-330.680
Maximum Log-likelihood	-243.621	-258.033	-195.779
Adjusted Log-likelihood ratio	0.255	0.216	0.398
Number of samples	301	301	301

(t-statistics in parentheses; *: significant at 5 %; **: at 1%)

In order to analyze the influence of time length of alternatives, we compare the results of Case 1 and 2. Commonly, the t-value of σ_{23}, which indicate the similarity between the alternatives 2 and 3, is significant. These statistics seem to stem from that the responded office arrival time was concentrated close to the boundary between the alternatives 2 and 3. If the definition of sequential time categories is shifted by a few minutes later, many respondents will move to another sequential alternative. This is a reason why the similarity between two alternatives highly occurred. Hence it is required to employ the PCL model dealing with the unavoidable similarity among alternatives

In order to analyze the influence of different time period definitions, we compare the estimation results of Cases 1 and 3. The comparison of the parameter estimates and t-statistics indicates that higher similarity would occur when the boundary of categories is set up to close to the peak of the distribution of responded time. The similarity parameter σ_{23} is significantly higher in Case 1, but not in Case 3.

The alternatives 1 and 3 are not sequential, nevertheless the similarity parameters σ_{13} are significant in all cases. This implies that there exist common unobserved variables, such as preference and taste of activity time at home or office, in these two alternatives.

The scale parameters of the alternatives 3 are restricted to 1.0 in estimating HEV model. In Cases 1 and 2, the estimated scale parameters θ_1 of the alternative 1 are larger than 1.0, while those of the alternative 2 are smaller than 1.0. This means that travelers who choose the earliest time alternative are apt to change their choices more randomly. Besides, the estimated scale parameters of θ_1 and θ_2 in Case 1 are higher than those in Case 2. The error variance of the alternative 2 with wider length of time becomes larger. It is found that many respondents concentrate on both boundaries of the time alternative 2. Consequently, the error between the estimated and actual choices arises more frequently under such definition. The estimated scale parameters in Case 3 show that θ_2 is the largest and then θ_1 follows. Hence, this reflects the respondents tend to concentrate in the middle of time choice alternatives in Case 3.

These results suggest that the heteroscedasticity among alternatives clearly exists in time choice behavior and is affected by the time categories.

To sum up, if continuous time is categorized near the peak of distribution of observed time choices, the similarity and heteroscedasticity cannot be ignored in this case study. The optimal predetermination of the time categories is primarily required to avoid such problems. However, there still remain some non-IID elements after the optimal categories. From this point of view, the application of flexible models, like the PCL and HEV models, is effective in averting the biases in the parameter estimates of conventional MNL model.

5. CONCLUSION

We summarized the non-IIA models that can relax the IID assumption fully or partially. Two models considering the similarity or heteroscedasticity of alternatives are applied to time choice behavior. The remarkable findings obtained from this study are as follows:

1. The IID assumption of MNL models is not maintained with the context of time choice because the similarity and heteroscedasticity appear among time choice alternatives.

2. By comparing various cases of time periods and time lengths of alternatives, it is confirmed that the degrees of similarity and heteroscedasticity variant over these cases.

3. The concentration of time choices to the boundaries of the predetermined time alternatives causes the stronger similarity and heteroscedasticity.

However, it seems impossible for analysts to avoid the similarity and heteroscedasticity by optimizing the time categories prior to model buildings. Applying non-IIA models, like the PCL or HEV model, will contribute to acquitting such burdensome problem and will enable to estimate with less biases in the parameter estimates.

Table 7. Estimation results of HEV model

Variable	Case 1	Case 2	Case 3
Category characteristics			
Time boundary between Alt. 2 and 3	9:00	9:00	9:30
Time range of Alt. 2	60 min	40 min	60 min
Commuting travel time	-0.104	-1.012	-0.348
(hour)	(-0.144)	(-0.931)	(-1.154)
Commuter by car	0.101	0.632	0.165
(dummy)	(0.144)	(0.856)	(0.485)
Commuter by public transportation	0.047	-2.212	0.553
(dummy)	(0.137)	(-0.552)	(0.434)
Age	0.007	0.102	0.047
(year)	(0.155)	(1.466)	(2.303)*
Single	0.032	0.565	-0.371
(dummy)	(0.133)	(0.521)	(-0.768)
Constant for Alt. 2	3.226	5.543	1.829
	(1.500)	(2.124)*	(1.609)*
Constant for Alt. 3	3.091	5.603	3.386
	(1.781)	(1.785)	(2.974)**
Scale parameter (θ_3 is fixed to 1.0)			
θ_1	4.167	1.298	5.180
θ_2	0.567	0.376	9.691
Initial Log-likelihood	-330.680	-330.680	-330.680
Maximum Log-likelihood	-243.314	-258.010	-195.575
Adjusted Log-likelihood ratio	0.257	0.217	0.399
Number of samples	301	301	301

(t-statistics in parentheses; *: significant at 5 %; **: at 1%)

REFERENCES

a) Books and Books chapters

Ben-Akiva and M.,Lerman, S. (1985) **Discrete Choice Analysis,** MIT Press.

Hensher, D. and Button, K. (2000) **Handbook of Transportation Modeling,** Pergamon.

Little, R. and Rubin, D. (1987) **Statistical Analysis with Missing Data**, John Wiley and Sons, New York.

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

b) Journal papers

Bhat, C. (1995) A Heteroscedastic Extreme-Value Model of Intercity Mode Choice. **Transportation Research, 29B, No. 6,** 471-483.

Bhat, C. (1998) Accommodating Variations in Responsiveness to level-of-service

Brownstone, D. and Train, K. (1999) Forecasting New Product Penetration with Flexible Substitution Patterns. **Journal of Econometrics, Vol. 89,** 109-129.

Dempster, A., Laird, N. and Rubin, D. (1977) Maximum likelihood from incomplete data via the EM algorithm, **Journal of the Royal Statistical Society B 39**, 1-38.

Fukuda, D. and Morichi, S. (1999) Review of Arrangement of Alternatives in Discrete Choice Models and Comparative Studies in the Recreational Destination Choice Behavior. **Journal**

of the Eastern Asia Society for Transportation Research, Vol. 3, No.5, 229-243

Koppelman F. and Wen C. (2000) The Paired Combinatorial Logit Model: Properties, Estimation and Application. **Transportation Research, Part B, Vol. 34**, 75-89.

McFadden, D. (1978) Modeling the Choice of Residential Location. **In Transportation Research Record 672, TRB**, National Research Council, Washington D.C., 72-77.

Recker, W. (1995) Discrete Choice with an Oddball Alternative. **Transportation Research, 29B**, 201-212.

Sugie, Y. Zhang, J. and Fujiwara, A. (1999) Dynamic Discrete Choice Models Considering Unobserved Heterogeneity with Mass Point Approach. **Journal of the Eastern Asia Society for Transportation Research, Vol. 3, No.5,** 245-260

Suto, K., Sugie, Y., Fujiwara, A. (1998) Changes in Commuting Travel Behavior after Introducing Flextime System, **Infrastructure Planning review, No. 15**, 655-662

Yai, T., Iwakura, S. and Morichi, S. (1997) Multinomial Probit with Structured Covariance for Route Choice Behavior. **Transportation Research, B 31,** 195-207.

c) Papers presented to conferences

Bhat, C. (1997) Recent Methodological Advances Relevant to Activity and Travel Behavior Analysis. **The 8th Meeting of the International Association for Travel Behavior Research**, Austin, Texas.

Brownstone, D. (2000) Discrete Choice Modeling for Transportation, **The 9th Meeting of the International Association for Travel Behavior Research**, Australia

Chu, C. (1989) A Paired Combinatorial Logit Model for Travel Demand Analysis. **Proceedings of the Fifth World Conference on Transportation Research, Vol. 4,** Ventura, CA, 295-309.

McFadden, D. (2000) Disaggregate Behavioral Travel Demand's RUM Side, A 30-Year Retrospective. **The 9th Meeting of the International Association for Travel Behavior Research**, Australia

Srinivasan, K. and Mahamassani, H. S., (2000) Kernel Logit Method for the Longitudinal Analysis of Discrete Choice Data: Some Numerical Experiments. **The 9th Meeting of the International Association for Travel Behavior Research**, Australia

Vovsha, P. (1997) The Cross-Nested Logit Model: Application to Mode Choice in the Tel-Aviv Metropolitan Area. **Presented at the 76th Transportation Research Board Annual Meeting,** Washington D.C., 1997

Wen C. and Koppelman F. (2000) The Generalized Nested Logit Model. **Presented at the 79th Transportation Research Board Annual Meeting,** Washington D.C., 2000

d) Other documents

Daganzo, C. (1979) Multinomial Probit: The Theory and Its Application to Demand Forecasting. **Academic Press**, New York.

Econometric Software (1994) **LIMDEP**. Econometric Software Inc., New York and Sydney.

McFadden, D. and Train, K. (1997) Mixed MNL models of Discrete Choice. **Working Paper**, Dept. of Econometrics, University of California, Berkeley