Modeling Heterogeneous Dynamics in the Stated and Revealed Travel Mode Choice Behaviors with Panel Data[*]

by Zhuo WANG[**], Junyi ZHANG[***] and Akimasa FUJIWARA[****]

1. Introduction

Travel behavior changes over time in both long-term and short-term perspective, and such changes might be not the same across individuals, either. The factors influence individuals travel behavior changes include many aspects, such as individuals' choice history, household or workplace attributes, the circumstantial context, the alternatives' attributes, and so on. This paper deals with travel mode choice by especially focusing on individual tastes to levels of services (e.g., travel time, cost and frequency), which are the most important factors to shape transportation planning and demand management policies. By assuming that change of taste per unit of time, "ct", is stable, but might show changing influences on travel behavior due to the lapse of time, individual taste at time t, B(t), can be defined as the summation of the taste at initial time, B(0), and the accumulated influences of "ct" over time, which are allowed to further differ across individuals. Conventional dynamic discrete choice models have not satisfactorily represented individual heterogeneity due to the insufficient utilization of individual attributes and behavior data. This paper attempts to incorporate such heterogeneous dynamics into the dynamic GEV (DGEV) modeling framework proposed by Swait and his colleagues [1)]. The DGEV model can simultaneously represent initial conditions, state dependence and influence of future expectation. An empirical analysis is conducted using a 4-wave panel data collected in Hiroshima city from 1987 to 1994. The data include revealed preference (RP) panel survey data and stated preference (SP) panel survey data, the estimated results confirm that the established model can flexibly describe choice behavior dynamics with a higher accuracy using both RP and SP data.

2. Model

It is known that Heckman [2)] presents a general modeling framework of dynamic behavior, as shown below.

$$u_{ijt} = v_{ijt} + \varepsilon_{ijt} \tag{1}$$

$$v_{ijt} = \beta x_{ijt} + \sum_{k=1}^{\infty} \gamma_{t-k,t} d_{ij,t-k} + \sum_{k=1}^{\infty} \lambda_{k,t-k} \prod_{q=1}^{k} d_{ij,t-q} + G(L) u_{ijt} \tag{2}$$

Here, *i*, *j* and *t* indicate individual, alternative in choice set and time, respectively. u_{ijt} is the utility function, v_{ijt} is deterministic portion and ε_{ijt} is the error component. d_{ijt} is a variable indicate choice result, equal to 1 when $u_{ijt} \geq 0$ and 0 when $u_{ijt} < 0$. $\gamma_{t-k,t}$ explains the influence of previous choice results on current choice, i.e., true state dependence. $\lambda_{k,t-k}$ describes the accumulated effects of previous choice results. x_{ijt} is explanatory variable with parameter β. $G(L)$ is a lag operator that represents the influence of past preference (i.e., behavior inertia).

Heckman's dynamic model is very general and can include many existing models as special cases. Recently, in line with the idea of Heckman's model, Swait et al.[1)] derived a new dynamic model by specifying the following *G* function of the well-known GEV model family.

$$G(y_{ijt}) = \sum_{j} \left\{ \prod_{s=1}^{\infty} \gamma_{ijs} y_{ijt+s} \cdot \prod_{s=0}^{t} \alpha_{ijs} y_{ijt-s} \right\}^{\mu_t} \tag{3}$$

$$y_{ijt} = \exp(v_{ijt}) \tag{4}$$

---

[*] Keywords: Travel mode choice, dynamic analysis, heterogeneity and DGEV model

[**] Member of JSCE, M. Eng, Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1Kagamiyama, Higashi-Hiroshima, 739-8529, Japan, TEL&FAX 082-424-6919; wangzhuo918@hiroshima-u.ac.jp

[***] Member of JSCE, Dr. Eng, Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1Kagamiyama, Higashi-Hiroshima, 739-8529, Japan, TEL&FAX 082-424-6919; zjy@hiroshima-u.ac.jp

[****] Member of JSCE, Dr. Eng, Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1Kagamiyama, Higashi-Hiroshima, 739-8529, Japan, TEL&FAX 082-424-6919; afujiw@hiroshima-u.ac.jp

where v_{ijt} is the deterministic term of utility function, γ_{ijs} is the parameter explaining future utilities, α_{ijs} describes the influence of past utilities (i.e., habit persistence and variety-seeking), and μ_t is scale factor at time *t*.

Under the principle of random utility maximization, a new dynamic GEV model (hereafter, called DGEV model) can be derived as:

$$P_{ijt} = \prod_{t=1}^{T} \frac{\exp(\widetilde{V}_{ijt})}{\sum_{j'} \exp(\widetilde{V}_{ij't})} \tag{5}$$

$$\widetilde{V}_{ijt} = \left(1 + \varphi_{ijt}\right) u_{ijt} + \sum_{s=1}^{t} \left(u_{ij,t-s} + \ln \alpha_{ijs}\right) \tag{6}$$

where $\widetilde{V}_{ijt}$ is the meta-utility, φ_{ijt} represents the influence of future expectation ($\varphi_{ijt} \geq 0$) and is transformed from γ_{ijs}, and T is the number of total time points. Using equation (6), initial condition, future behavior (expectation), state dependence, and time-varying taste can be simultaneously incorporated.

Omitting the 2nd to 4th items in the right side of Equation (2), the utility function can be represented below.

$$u_{ijt} = \beta_{it} x_{ijt} + \varepsilon_{ijt} \tag{7}$$

Since the temporally-changing taste parameter can be transformed into

$$\beta_{it} = \beta_{it-1} + \Delta\beta_{it} \tag{8}$$

then, the utility function can be decomposed into

$$u_{ijt} = \beta_{it-1} x_{ijt-1} + \beta_{it} \Delta x_{ijt} + \Delta\beta_{it} x_{ijt-1} + \varepsilon_{ijt} \tag{9}$$

where $\beta_{it-1} x_{ijt-1}$ is the previous utility, $\beta_{it} \Delta x_{ijt}$ is the change of level of variable x and $\Delta\beta_{it} x_{ijt-1}$ is the behavioral adjustment.

To derive an operational dynamic model for future prediction, it is necessary to explore the regularity in temporal change. Therefore, we propose to further transform taste parameter as follows:

$$\left.\begin{aligned} &\beta_{it} = \beta_{it-1} + \rho_{it} \Delta\beta, \qquad t=1,\ldots,T \\ &\beta_{it} = \beta_{i0} + \Delta\beta \sum_{s=1}^{t} \rho_{is} \\ &\rho_{it} = \tau^{a_{it}}, \quad a_{it} = kZ_{it} \end{aligned}\right\} \tag{10}$$

where β_{i0} is the taste parameter at the initial time point t_0, $\Delta\beta$ is the change of taste per unit, and ρ_{it} is the influence of taste due to the progress of time on behavior.

In order to compare the validity of the DGEV model with heterogeneity (called DGEV_H model), we also estimate a dynamic MNL model with heterogeneity, which is a special case of DGEV_H model without the consideration of future behavior, as shown below. Note β_{it} is defined using equation (10).

$$P_{ijt} = \prod_{t=1}^{T} \frac{\exp(\beta_{it} x_{ijt})}{\sum_{j'} \exp(\beta_{it} x_{ij't})} \tag{11}$$

3. Data and Model Estimation

The data using in this study was collected by our laboratory in Hiroshima city from 1987 to 1994. The purpose is to predict commuter demand for Astramline, opened at the time of the 12th Asian Game in 1994, our laboratory conducted a 4-wave SP panel survey (i.e., the years of 1987, 90, 93, 94). The SP panel survey considers passenger car and bus as the alternative modes of Astramline. Each respondent was requested to participate in the panel as much as possible. Sample refreshment was also done to cover for samples dropped out during the course of the panel survey.

In the survey conducted before the opening of Astramline, the respondents were asked to answer several hypothetical choice questions, meanwhile, report their actual choice modes for commuting. After its opening, these panel respondents were asked again to report their actual travel modes including Astramline. Although multiple SP choice results were obtained in the survey from most of the respondents, they are regarded as single-choice results from different respondents without loss of generality. As a result, 226 valid panel samples were obtained.

In this model, gender, occupation and number of household members were chosen as explanatory variables, as well as the level of service variables, which are not the same between RP and SP data (see Table 1). The data used in this study covers 4 points in time, i.e., 1987, 1990, 1993and 1994. We estimate the DGEV_H and DMNL_H models for six time periods, i.e., 87-90, 87-90-93, 87-90-93-94, 90-93, 90-93-94, and 93-94. Due to the limited spaces, for comparison of model accuracy, Table 2 only shows the results of one model for each of SP and RP data, respectively. It is clear that for both RP and SP data, DGEV_H models are clearly superior to DMNL_H model.

Table 1. Explanatory variables

RP Data		SP Data
travel time, travel cost (including parking fee)	CAR	travel time, travel cost (including parking fee)
travel time, travel cost	BUS	travel time, travel cost, waiting time at bus stop
	NTS Astramline	travel time, travel cost, access time to and waiting time at Astramline station

Table 2. Comparison of model accuracies

Adjusted McFadden's Rho-squared	DMNL_H model	DGEV_H model
RP Data (1987, 1990, 1993, 1994)	0.3678	0.4596 (25% improved)
SP Data (1987, 1990, 1993, 1994)	0.1079	0.2898 (169% improved)

Table 3 shows the estimation results of DGEV_H models from 1987 to 1994 for both RP and SP data. McFadden's Rho-squared is 0.4937 and 0.3157 for RP and SP model respectively, suggesting that the model accuracy is satisfactory. Observing all of the parameters, except for the parameters of access time, more than half of the parameters are statistically significant larger 95% level. Initial utility and state dependence estimate by different models are shown in Fig 1 (RP) and Fig 2 (SP). All of state dependence parameters are positive, it is understood that individuals prefer to maintain their habit rather than pursue the diversity when choosing travel mode. Contrarily, the parameters of initial utility present instability in this model estimate results, this means it needs the further study in the future and to find a more suitable variable. Fig 3 and Fig 4 show the estimated results of influence of end-time of panel survey based on future expectation. It can be seen from the figures, if the start-time of survey is fixed and the end-time is changed then the estimated future expectation parameters will show different change pattern. For instance, when the survey start-time is fixed in 1987, and the end-time is changed in 1990, 1993, 1994, the change patterns of the future expectation parameters are not steady. It is thought that the number of time-points will influence the estimated results, but compare the results with the same number of time-points (i.e., 87-90, 90-93, 93-94 with two time-points and 87-90-93, 90-93-94 with three time-points), it is still there is not a stable change pattern of the future expectation parameter. Moreover, the value of the future expectation parameter with different pattern when the end-time in 1993 and 1994 respectively; at the same time, the value of the future expectation parameter become smaller while approaching in 1994. This can be interpreted that the influence of future expectation be reduced over time. Therefore, it is suggested that model estimations might be different according to end-time of the panel survey.

4. Conclusion

It is confirmed that the developed DGEV_H model is effective to capture heterogeneous dynamics in travel mode choice behavior in both RP and SP contexts. State dependence (habit persistence) seems not sensitive to time frame of panel survey. However, end-time of panel survey surely influences other aspects of travel behavior, i.e., decision of terminating a panel survey should be paid much more sufficient attention.

References

1) Swait, J., Adamowicz, W. and van Bueren, M: Choice and temporal welfare impacts: Incorporating history into discrete choice models. Journal of Environmental Economics and Management, 47, 94-116, 2004.

2) Heckman J.J: Statistical models for discrete panel data. In: Manski C.F., McFadden D. (Eds.), The Structural Analysis of Discrete Data, Cambridge: MIT Press, 114-178, 1981.

Table 3. DGEV_H models: RP and SP

Explanatory variables	RP Model	SP Model
Constant term of BUS		
Initial value	-1.4047**	-0.3640
Average annual change (Δβ)	0.0631**	0.4122**
Influence of 1987 year on average annual change	0.1440	0.5288
Influence of time lapse on average annual change		
Constant term	-0.5167	-1.5646
Gender	0.1528	1.1042**
Age	0.0475*	-0.0429**
Occupation	-0.8509	1.7787*
Number of household	-0.7779*	0.2150**
Travel cost		
Initial value	-0.0029**	-0.0017**
Average annual change (Δβ)	0.0005**	0.0043**
Influence of 1987 year on average annual change	2.3668**	0.3065**
Influence of time lapse on average annual change		
Constant term	-1.7107*	-0.1075
Gender	-0.8171**	1.2476**
Age	0.0325*	-0.0542**
Occupation	0.7506	0.4201
Number of household	-0.0211	0.3550**
Travel time		
Initial value	-0.0114	0.0019
Average annual change (Δβ)	-0.0238**	-0.0038*
Influence of 1987 year on average annual change	-0.7580*	0.1178
Influence of time lapse on average annual change		
Constant term	-0.8690	0.1165
Gender	0.5206	1.0527**
Age	-0.0194	-0.0062
Occupation	1.1698	-2.0381**
Number of household	-0.1452*	-0.4082*
Constant term of NTS		
Initial value		0.3605
Average annual change (Δβ)		0.0012
Influence of 1987 year on average annual change		1.9988*
Influence of time lapse on average annual change		
Constant term		-0.8418
Gender		1.4842**
Age		-0.1330**
Occupation		0.5688
Number of household		1.9440**
Waiting time at bus stop or NTS station		
Initial value		0.2875**
Average annual change (Δβ)		0.5798**
Influence of 1987 year on average annual change		-0.4850**
Influence of time lapse on average annual change		
Constant term		0.1877
Gender		-0.5682
Age		-0.0223
Occupation		-0.7221
Number of household		-1.9021**
Access time to NTS station		
Initial value		-0.0097
Average annual change (Δβ)		0.0017
Influence of 1987 year on average annual change		0.1924
Influence of time lapse on average annual change		
Constant term		1.0471
Gender		0.0271
Age		-0.0078
Occupation		0.1518
Number of household		-0.1530**
Model goodness-of-fit		
Initial log-likelihood	-676.99	-1256.20
Converged log-likelihood	-342.73	-859.64
McFadden's Rho-squared	0.4937	0.3157
Adjusted McFadden's Rho-squared	0.4727	0.2618
Sample size	904=226*4	904=226*4

Note: * significant at 95% level; ** significant at 99% level

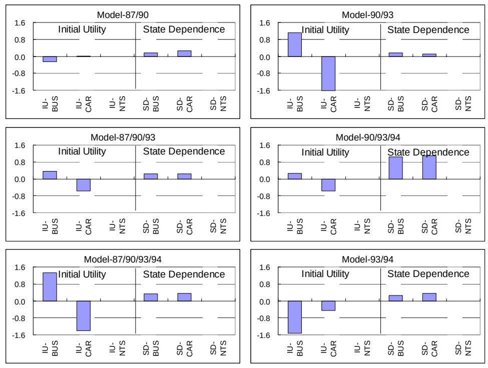

Fig 1 Initial utility and state dependence: RP (DGEV_H model)

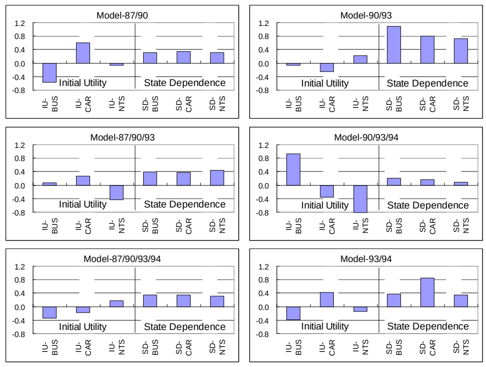

Fig 2 Initial utility and state dependence: SP (DGEV_H model)

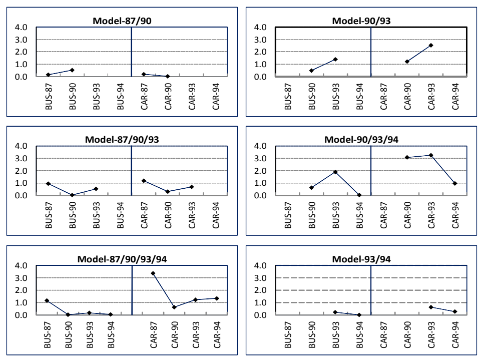

Fig 3 Influence of end-time of panel survey based on future expectation: RP

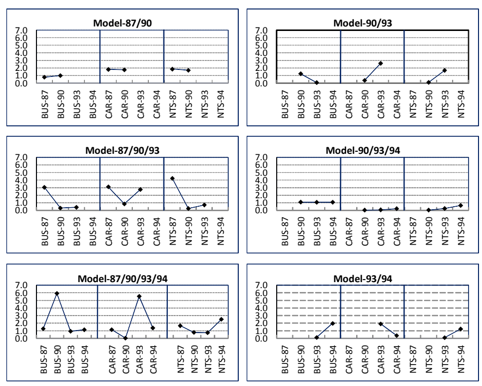

Fig 4 Influence of end-time of panel survey based on future expectation: SP