Journal of the Eastern Asia Society for Transportation Studies, Vol.8, 2010

Representing the Dynamics in Stated Travel Choice Behavior Based on a DGEV Model with Heterogeneity

Zhuo WANG
Doctoral Course Student
Graduate School for International Development and Cooperation
Hiroshima University
1-5-1, Kagamiyama, Higashi-Hiroshima
739-8529 Japan
Fax: +81-82-424-6919
E-mail: wangzhuo918@hiroshima-u.ac.jp

Junyi ZHANG
Associate Professor
Graduate School for International Development and Cooperation
Hiroshima University
1-5-1, Kagamiyama, Higashi-Hiroshima
739-8529 Japan
Fax: +81-82-424-6919
E-mail: zjy@hiroshima-u.ac.jp

Akimasa FUJIWARA
Professor
Graduate School for International Development and Cooperation
Hiroshima University
1-5-1, Kagamiyama, Higashi-Hiroshima
739-8529 Japan
Fax: +81-82-424-6919
E-mail: afujiw@hiroshima-u.ac.jp

Abstract: Travel behavior changes over time and such changes might be not the same across individuals, either. In the context of stated choices, such behavioral dynamics might become more complicated. This paper proposes a new DGEV (Dynamic Generalized Extreme Value) model with heterogeneity for not only SP data, but also the combined SP/RP data to further improve the predictability of the stated preference (SP) data. The DGEV model is used to simultaneously capture initial conditions, state dependence, and future expectation as well as time-varying tastes. Heterogeneity is measured with respect to the observed individual tastes to levels of travel services. An empirical analysis is conducted using a 4-wave panel data of travel mode choice collected in Hiroshima City in 1987, 1990, 1993 and 1994, where a new transit system was targeted. It is confirmed that the proposed models are effective to capture heterogeneous dynamic in travel mode choice behavior.

Key Words: *revealed preference, stated preference, DGEV model*

1. INTRODUCTION

Transportation planners need to understand people's responses in case of introducing some new policies, which are however sometimes difficult to be evaluated using RP (revealed preference) data. One can apply RP data to evaluate the above-mentioned new policies; however, it has to be assumed that RP model could be applicable to SP situations. Such assumption might be unrealistic. This is because RP data only include the information about existing alternatives in choice set and might not cover the variation information of people's responses to the new policies. To overcome the shortcomings of RP approach, SP (stated preference) approach was proposed (Louviere *et al.* 2000). SP approach examines individual responses to a series of experimentally designed choice alternatives, which are typically described in terms of combinations of attributes with several pre-defined levels. SP data are particularly rich in attribute tradeoff information because wider attribute ranges can be built

into experiments, which in turn, allow models estimated from SP data to be more robust than models estimated from RP data (Swait *et al*. 1994). Because of such experimental characteristics, SP approach is usually used to predict people's responses under not-yet-existing alternatives. In this sense, SP data can cover a much wider range of attributes and levels than RP data. Besides the ability to directly measure the demand/response under not-yet-existing conditions, SP approach is also able to control statistical problems such as multi-collinearity and lack of variance in explanatory variables, to include qualitative factors as explanatory variables, and it is more cost-efficient to develop models from a relatively small size of samples than RP approach.

Because of the consideration of hypothetical situations in SP approach, respondents are sometimes placed in unfamiliar situations in which complete information is not available (Whitehead *et al*. 2007). As a result, the SP approach inevitably includes some biases such as reporting bias and non-reporting bias. SP reporting bias arises when the respondents do not answer the questions based on the attributes presented in profiles or misunderstand/ignore the trade-off relationships among these attributes. SP non-reporting bias occurs when the respondents do not give answers about the questions. Another bias is called fatigue bias caused by repeatedly answering the SP questions. To relax some of the SP biases, the combined SP/RP model were developed in the early 1990s (Ben-Akiva and Morikawa, 1990 and Morikawa, 1994; Bradley and Daly, 1997). The basic concept of this combined model is to correct the SP biases by using the RP information. Concretely speaking, SP and RP data are combined together by introducing a scale parameter into the SP utility function and at the same time assuming the same parameter for the same explanatory variable in RP and SP utility functions. Needless to say, SP and RP utility functions could have some variables different from each other. The scale parameter is introduced to balance the different variances of SP and RP utility functions, which are expected to be caused by the different reliabilities in these two data sources. On the other hand, travel behavior changes over time and such changes might be not the same across individuals, either. In the context of stated choices, such behavioral dynamics might become more complicated. Under such circumstance, this paper proposes a new combined SP/RP DGEV model. Different from existing studies, this study develops the combined model in the DGEV (Dynamic Generalized Extreme Value) modeling framework to further improve the predictability of the SP data. DGEV model, proposed by Swait *et al*. (2004), is used to simultaneously capture state dependence and future expectation as well as initial conditions. Heterogeneity is measured with respect to the observed individual tastes to levels of travel services.

As an empirical study, a 4-wave panel data, collected in Hiroshima City, Japan in 1987, 1990, 1993 and 1994, are used in this study. The data include both RP and SP information related to commuting travel mode choice behavior (alternatives are car, bus and a new transit system). For comparison, a multinomial logit (MNL) model is established with respect to RP and SP data, respectively, where the former is called RP model and the latter SP model. In addition, a traditional combined SP/RP model is also estimated. Therefore, this paper is organized as follows: the next section introduces the development of the DGEV model and the method for combining RP and SP data. Section 3 describes the data, estimates and discusses the models. Finally, this study is concluded in Section 4.

2. MODEL

Since this study deals with dynamic discrete travel choice behavior, Heckman's dynamic model (Heckman, 1981) is relevant. Heckman developed a general structure for discrete choice models as follows:

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

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

where,

n, i, t	: individual, alternative in choice set and time, respectively,
u_{nit}	: utility function that individual n chooses alternative i at time t,
v_{nit}	: deterministic term of utility function u_{nit},
ε_{nit}	: error term of utility function u_{nit},
d_{nit}	: choice result that is equal to 1 when chosen, otherwise 0,
$\gamma_{t-k,t}$	: true state dependence, indicating the influence of previous choice results on current choice,
$\lambda_{k,t-k}$	: the accumulated effects of previous choice results on current choice,
x_{nit}	: explanatory variable with taste parameter β_{nt}, and
$G(L)$	: lag operator, representing the influence of past preference on current choice (i.e., behavior inertia).

Heckman's dynamic model can include many existing models as special cases. Recently, in line with the idea of Heckman's model, Swait *et al*. (2004) derived a new dynamic model by specifying the following G function of the well-known GEV model family.

$$G(y_{nit}) = \sum_{i} \left\{ \prod_{s=1}^{\infty} \gamma_{nis} y_{ni,t+s} \cdot \prod_{s=0}^{t} \alpha_{nis} y_{ni,t-s} \right\}^{\mu_t} \tag{3}$$

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

where,

γ_{nis}	: parameter explaining future utilities (i.e., future expectation),
α_{nis}	: influence of past utilities (i.e., habit persistence and variety-seeking), and
μ_t	: scale factor at time t.

Transforming equations (3) and (4) under the principle of random utility maximization, Swait *et al.* successfully derived a new dynamic GEV model (called DGEV model), as shown below.

$$P_{nit} = \frac{exp(\tilde{V}_{nit})}{\sum_{i'} exp(\tilde{V}_{ni't})} \tag{5}$$

$$\tilde{V}_{nit} = (1 + \varphi_{nit}) v_{nit} + \sum_{s=1}^{t} (v_{ni,t-s} + \ln \alpha_{nis}) \tag{6}$$

where,

$\tilde{V}_{nit}$	: deterministic term of meta-utility, and

φ_{nit} : influence of future expectation ($\varphi_{nit} \geq 0$).

Equation (6) can simultaneously represent initial condition, future expectation, state dependence, and time-varying taste, etc. Note that v_{nit} is defined in equation (2).

Omitting the 2nd to 4th items in the right side of equation (2), the utility function u_{nit} can be represented below.

$$u_{nit} = v_{nit} + \varepsilon_{nit} = \beta_{nt} x_{nit} + \varepsilon_{nit} \tag{7}$$

Since the temporally-changing taste parameter β_{nt} can be naturally transformed into

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

where, β_{nt-1} is the taste parameter of individual *n* at time point *t-1*, and $\Delta\beta_{nt}$ is the variation of taste parameter at time point *t*.

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

$$\beta_{nt} = \beta_{nt-1} + \rho_{nt}\Delta\beta \tag{9}$$

where,

$\Delta\beta$: the change of taste parameter per unit of time, and

ρ_{nt} : the influence of taste parameter due to the progress of time on behavior.

Let β_{n0} be the taste parameter at the initial time point t_0, then equation (9) can be re-written as

$$\beta_{nt} = \beta_{n0} + \Delta\beta \sum_{s=1}^{t} \rho_{ns} \tag{10}$$

$$\rho_{nt} = \tau^{a_{nt}} \ , \ a_{nt} = \pi Z_{nt} \tag{11}$$

where,

a_{nt} : composite variable of individual attributes at time point *t*, and

Z_{nt} : vector of individual attributes at time point *t* with parameter vector π.

Equations (10) and (11) are used in this study to represent the influence of observed heterogeneity on choice behavior. The observed heterogeneity is named because all the individual attributes Z_{nt} are observed.

Substituting equations (10) and (11) results in the DGEV model with heterogeneity (called DGEV-H model). For the purpose of comparison, the following multinomial logit model with equation (7) is first established (called *Temporally-Changing Parameter* (TCP) model), which assumes that all the taste parameters are invariant across individuals. Using TCP model, we can calculate the individual taste parameter at each time point.

$$P_{nit} = \frac{exp(v_{nit})}{\sum_{i'} exp(v_{ni't})} = \frac{exp(\beta_{nt} x_{nit})}{\sum_{i'} exp(\beta_{nt} x_{ni't})} \tag{12}$$

Another comparison model is built below by using equations (7), (10) and (11) in the multinomial logit modeling framework. Since in equations (10) and (11), the heterogeneity is specified with respect to both individual and time, the established model is called *Cross-sectional and Longitudinal Heterogeneity* (DCLH) model.

$$P_{nit} = \frac{exp(v_{nit})}{\sum_{i'} exp(v_{ni't})} = \frac{exp((\beta_{n0} + \Delta\beta \sum_{s=1}^{t} \rho_{ns}) x_{nit})}{\sum_{i'} exp((\beta_{n0} + \Delta\beta \sum_{s=1}^{t} \rho_{ns}) x_{ni't})} \tag{13}$$

If individual attributes (Z_{nt}) are not introduced into the DCLH model, the resulting model is called homogeneous DCLH (simply Ho-DCLH) model, as shown below, where ρ_{nt} in equations (9) ~ (11) becomes ρ_t (or ρ_s). This model will be also used to compare with DGEV-H model. Here, homogeneity means that the influence of taste parameter due to the progress of time on behavior does not change across individuals.

$$P_{nit} = \frac{exp(v_{nit})}{\sum_{i'} exp(v_{ni't})} = \frac{exp((\beta_{n0} + \Delta\beta \sum_{s=1}^{t} \rho_{s}) x_{nit})}{\sum_{i'} exp((\beta_{n0} + \Delta\beta \sum_{s=1}^{t} \rho_{s}) x_{ni't})} \tag{14}$$

In summary, the above-established four models, i.e., TCP, DCLH, Ho-DCLH and DGEV-H, are used to predict the commuting demand of a new transit system in this study. These four models are further estimated with respect to SP data and the combined SP and RP data, respectively.

In case of using the combined SP/RP model, the utility functions for RP and SP data are written as follows:

$$u_{nit}^{rp} = v_{nit}^{rp} + \varepsilon_{nit}^{rp} = \beta x_{nit}^{rp} + \alpha w_{nit}^{rp} + \varepsilon_{nit}^{rp} \tag{15}$$

$$u_{nit}^{sp} = v_{nit}^{sp} + \varepsilon_{nit}^{sp} = \beta x_{nit}^{sp} + \gamma w_{nit}^{sp} + \varepsilon_{nit}^{sp} \tag{16}$$

where,

$x_{nit}^{rp}, x_{nit}^{sp}$: alternative-specific variables that are common in both RP and SP utilities,

$w_{nit}^{rp}, w_{nit}^{sp}$: alternative-specific variables that are different for each type of data, and

$\varepsilon_{nit}^{rp}, \varepsilon_{nit}^{sp}$: alternative-specific variables that are different for each type of data, and

α, β, γ : parameters to be estimated.

The probability function for the combined SP/RP model can be formulated as,

$$P_{nit} = P_{nit}^{rp} \cdot P_{nit}^{sp} = \frac{exp(v_{nit}^{rp})}{\sum_{i'} exp(v_{ni't}^{rp})} \cdot \frac{exp(\mu v_{nit}^{sp})}{\sum_{j} exp(\mu v_{ni't}^{sp})} \tag{17}$$

where, μ is the scale parameter of SP model, which is measured by assuming the scale

parameter of RP model to be equal to 1. μ has positive value and it is usually less than 1 because the variance of SP error term is usually larger than that of RP error term.

Then, the likelihood function for the dynamic combined SP/RP model can be described below, where T is the number of total time points.

$$L = \prod_{t=1}^{T} \left\{ \frac{exp(v_{nit}^{rp})}{\sum_{i'} exp(v_{ni't}^{rp})} \cdot \frac{exp(\mu v_{nit}^{sp})}{\sum_{j} exp(\mu v_{ni't}^{sp})} \right\} \tag{18}$$

Substituting the deterministic terms of the above TCP, DCLH, Ho-DCLH and DGEV-H models leads to the SP/RP_TCP, SP/RP_DCLH, SP/RP_Ho-DCLH and SP/RP_DGEV-H models. Maximum likelihood method will be used to estimates these models using TSP software.

3. DATA AND MODEL ESTIMATION

3.1 Data

The data using in this study was collected by our laboratory in Hiroshima City in 1987, 1990, 1993 and 1994. The purpose was to predict the commuting travel demand for a new transit system (NTS), called Astramline, opened at the time of the 12th Asian Game in 1994. The survey is a 4-wave panel survey for both RP and SP data. The RP data includes only car and bus as the alternative modes, and the SP panel data considers 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 from the respondents, they are regarded as single-choice results from different respondents without loss of generality. As a result, the valid data from 226 respondents were obtained.

Individual attributes of the respondents in this survey are summarized in Table 1. Note that only the information in the first year of the panel survey is shown. It is observed that 68.6% of the respondents are male, this might be because this survey selected the commuters using Astramline as the respondents. For the male respondents, most of them are aged between 30 and 60 years old. As for the female respondents, all of them are younger than 50 years old. And, 97.3% of the respondents are employed, and the remaining are school students. It is also revealed that 72.2% of the respondents belong to a family with 2 or more members.

Table 2 shows that in the RP survey the shares of bus users in the 4 waves are higher than those of car users, and there is an increasing trend of bus usage over time. However, in the SP survey, across the 4 waves, 45.1%~56.6% of respondents choose Astramline, but its share gradually decreases over time: the share is 56.6% in the first wave and decreases by more than 10% in the fourth wave. In contrast, bus share shows an increasing trend.

Table 1 Individual attributes of respondents (base on 1987)

Gender and age	Man (155 persons)	1: Under 30 years old	0	(0.0%)
		2: 30-39 years old	46	(29.7%)
		3: 40-49 years old	45	(29.0%)
		4: 50-59 years old	51	(32.9%)
		5: Over 60 years old	13	(8.4%)
	Woman (71 persons)	1: Under 30 years old	21	(29.6%)
		2: 30-39 years old	28	(39.4%)
		3: 40-49 years old	22	(31.0%)
		4: 50-59 years old	0	(0.0%)
		5: Over 60 years old	0	(0.0%)
Occupation status	Occupational		200	(97.3%)
	Non-occupational or Students		6	(2.7%)
Number of household	1 family members		63	(27.8%)
	2 family members		82	(36.3%)
	3 family members		51	(22.6%)
	4 family members		30	(13.3%)

Table 2 RP and SP mode choice

RP survey	1987	CAR	100	(44.2%)
		BUS	126	(55.8%)
	1990	CAR	94	(41.6%)
		BUS	132	(58.4%)
	1993	CAR	92	(40.7%)
		BUS	134	(59.3%)
	1994	CAR	92	(40.7%)
		BUS	134	(59.3%)
SP survey	1987	CAR	59	(26.1%)
		BUS	39	(17.3%)
		Astramline	128	(56.6%)
	1990	CAR	56	(24.8%)
		BUS	54	(23.9%)
		Astramline	116	(51.3%)
	1993	CAR	52	(23.0%)
		BUS	66	(29.2%)
		Astramline	108	(47.8%)
	1994	CAR	52	(23.0%)
		BUS	72	(31.9%)
		Astramline	102	(45.1%)

3.2 Model Estimation and Discussion

In the DCLH and DGEV-H models, gender, occupation and number of household members were chosen as explanatory variables to represent the heterogeneity, as well as the level-of-service variables, which are not the same between RP and SP data (see Table 3). We estimate all the models for all the time periods, i.e., 1987, 1990, 1993 and 1994 with respect to SP data and SP/RP combined data, respectively. Table 4 shows the model accuracies (McFadden's Rho-squared). For all of the four models with two types of data, the model with SP data shows the lowest accuracy except for the DGEV-H model. It is expected that such differences in model accuracy is caused by the different reliabilities of SP and RP data. This is also the reason why we attempt to make use of RP information to correct SP biases. It is clear that for the TCP, Ho-DCLH and DCLH models, the models with SP/RP combined data are clearly superior to those with only SP data (the improved rate of model accuracy ranges between 49% and

124%). Accordingly, it can be concluded that the model accuracies can be significantly improved by combining SP and RP data. The McFadden's Rho-squared of the DGEV-H model using SP/RP combined data is 0.2060, suggesting that model accuracy is satisfactory. But it is lower than the DGEV-H model with only SP data. We expected that the SP/RP_DGEV-H model should be better than any other models established in this study. Since the McFadden's Rho-squared of the SP/RP_DGEV-H model is 0.2060, we argue this model is acceptable. Since the DGEV-H model with only SP data shows the best model accuracy (0.3157), this implies that introducing behavioral dynamics with heterogeneity into the SP data under the DGEV modeling framework could also significantly correct the SP biases.

Table 3 Explanatory variables

RP Data	Travel Mode	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
	Astramline (NTS)	travel time, travel cost, access time to and waiting time at Astramline station

Table 4 Comparison of model accuracies

McFadden's Rho-squared	SP Data	RP+SP Data
TCP Model	0.0826	0.1786(116% improved than SP)
Ho-DCLH Model	0.0761	0.1708(124% improved than SP)
DCLH Model	0.1286	0.1910(49% improved than SP)
DGEV-H Model	0.3157	0.2060

Table 5 shows the estimation results of TCP model. When using only SP data, most of the parameters of travel cost and travel time are not statistically significant. But this does not directly mean that the stated choice behavior is stable over time, because the possible biases included in SP data have not been corrected. To clarify this point, the SP/RP combined model is estimated and it is observed that most of the above parameters become statistically significant. This implies that if the SP biases are not given a proper correction, travel behavior might be understood in a wrong way. The statistically significant scale parameter μ (located between the interval [0, 1]) also supports this conclusion. This suggests that people's responses to actual travel time and cost change over time. Comparing the SP and SP/RP models, it is found that there is no significant difference of t-score for waiting time and access time. This is because there is no RP information to correct these parameters. In the SP/RP model, most of the change parameters for travel time and travel cost in different time points are statistically significant. This suggests that people's responses to travel time and cost change over time.

Table 6 shows the estimation results of Ho-DCLH model. One remarkable difference is that the scale parameter (0.1633) in the SP/RP combined model in Table 6 is nearly twice larger than that (0.0931) in Table 5. Another difference is that the constant term of NTS in 1987 in Table 5 becomes much smaller than that in Table 6. Even though the Ho-DCLH cannot surely improve the model accuracy, it might improve the performances of model parameters.

Tables 7 shows the estimation results of DCLH model. Some of the statistically significant

parameters in the SP model become insignificant in the SP/RP combined model. This means

Table 5 Estimation results of TCP model

Explanatory variables	SP Data		RP+SP Data	
	Parameter	t-stats	Parameter	t-stats
Constants term of BUS				
Parameter of 1987 (β_t)	-0.3783	-0.6219	0.1823	1.0206
Change in 1990	0.1916	0.2680	0.0694	0.2914
Change in 1993	0.2366	0.2366	-0.2975	-1.1745
Change in 1994	-0.2930	-0.5459	0.4232	1.5779
Parameter of 1990 (β_t)	-0.1867	-0.4964	0.2518	1.5965
Parameter of 1993 (β_t)	0.0499	0.1333	-0.0457	-0.2305
Parameter of 1994 (β_t)	-0.2431	-0.6321	0.3775*	2.0891
Constants term of NTS				
Parameter of 1987 (β_t)	0.4692	0.7325	17.8205*	2.5121
Change in 1990	0.3470	0.3601	-11.2134	-1.4400
Change in 1993	0.1938	0.2016	-0.1000	-0.0130
Change i 1994	-1.1953	-1.2521	-8.0468	-0.9801
Parameter of 1990 (β_t)	0.8163	1.1335	6.6071	1.1140
Parameter of 1993 (β_t)	1.0101	1.5856	6.5071	1.0664
Parameter of 1994 (β_t)	-0.1852	-0.2604	-1.5397	-0.2890
Travel cost				
Parameter of 1987 (β_t)	-0.0021**	-3.5029	-0.0066**	-7.8529
Change in 1990	0.0020	1.8149	0.0031**	3.1259
Change in 1993	-0.0004	-0.3275	-0.0003	-0.4134
Change i 1994	0.0014	1.0404	0.0004	0.5153
Parameter of 1990 (β_t)	-0.0001	-0.1159	-0.0035**	-6.2256
Parameter of 1993 (β_t)	0.0005	-0.5679	-0.0038**	-6.6776
Parameter of 1994 (β_t)	0.0009	0.8972	-0.0034**	-6.8792
Travel time				
Parameter of 1987 (β_t)	-0.0409**	-2.9125	0.0253	1.4111
Change in 1990	0.0479**	3.0306	-0.0432*	-1.9866
Change in 1993	-0.0067	-0.6470	-0.0753**	-3.6365
Change i 1994	0.0011	0.0802	0.0477*	2.3051
Parameter of 1990 (β_t)	0.0070	0.9647	-0.0179	-1.4805
Parameter of 1993 (β_t)	0.0003	0.0349	-0.0932**	-5.4755
Parameter of 1994 (β_t)	0.0014	0.1176	-0.0455**	-3.8520
Waiting time at bus stop or NTS station				
Parameter of 1987 (β_t)	-0.0578	-1.1534	-0.4301	-1.6621
Change in 1990	0.1300	1.6774	0.7312	1.3966
Change in 1993	0.0216	0.2642	0.6278	0.9949
Change i 1994	0.0124	0.1573	0.1054	0.1815
Parameter of 1990 (β_t)	0.0723	1.2213	0.3011	0.7034
Parameter of 1993 (β_t)	0.0939	1.6622	0.9289	1.8219
Parameter of 1994 (β_t)	0.1063	1.9339	1.0344	1.9289
Access time to NTS station				
Parameter of 1987 (β_t)	-0.0784*	-2.0444	-0.6362	-1.4601
Change in 1990	0.0864	1.3120	0.6975	0.9572
Change in 1993	-0.0517	-0.6836	-0.5299	-0.6389
Change i 1994	0.0958	1.2676	0.9702	1.1102
Parameter of 1990 (β_t)	0.0080	0.1492	0.0613	0.1068
Parameter of 1993 (β_t)	-0.0437	-0.8181	-0.4686	-0.7914
Parameter of 1994 (β_t)	0.0521	0.9745	0.5015	0.8360
Scale parameter of SP model (μ)			0.0931**	2.9073
Initial log-likelihood	-993.1500		-1619.7506	
Converged log-likelihood	-911.1020		-1330.5400	
McFadden's Rho-squared	0.0826		0.1786	
Adjusted McFadden's Rho-squared	0.0744		0.1709	
Sample size	226*4=904		226*4=904	

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

that combining the SP and RP data does not necessarily improve the statistical performance of

all the parameters. Since the SP/RP model is superior to the SP model in theory, such insignificance should be acceptable. Focusing on the results of the SP/RP model, it is found that looking at the parameter π (i.e., the influence of individual attributes on temporal change

of taste), the individuals with different age and number of household members show differing responses to travel time over time. Such heterogeneity is also observed with respect to age and gender in the taste of waiting time. For the other variables, it seems that people do not show clearly heterogeneous responses. Concerning the scale parameter, it is smaller than that in Table 6, but larger than that in Table 5.

Table 8 shows the estimation results of the recommended model (i.e., DGEV-H model) in this study, which has the most complicated model structure. This model can simultaneously capture initial conditions, state dependence, and future expectation as well as time-varying tastes. Except for car in the SP model, other state dependence parameters are all statistically significant at the 99% level and positive. This can be interpreted that individuals prefer to maintain their habit rather than pursue the diversity when choosing travel mode. Contrarily, only the parameter of initial utility of car in the SP model is statistically significant and positive. This implies two things: one is that to properly describe car choice behavior, more information in the past should be introduced in to the model; another is that car choice behavior at a certain time does not influence the behavior in the short-term perspective (reflected by the insignificant state dependence parameter), but have the influence on the behavior in the long-term perspective. This is a very interesting finding, which can be only obtained using the DGEV model. Looking at the SP/RP model (i.e., SP/RP_DGEV-H model), travel cost parameter shows clearly temporal change and is heterogeneous with respect to occupation and number of household members as well as gender (the significance level is a little bit lower than others). Response of access time to NTS station also changes over time, and differs with respect to only occupation. Focusing on future expectation (φ_{nit}), for the existing travel modes of car and bus, the influence on the behavior in 1987 shows extremely higher values than those in other years, which are all not significant. Different from the existing travel modes, for the not-yet-existing travel mode, the influence of future expectation could last for a longer time in the sense that the parameter of the influence of future expectation on 1994 is statistically significant. In addition, all the parameters related to constant terms are not statistically significant.

4. CONCLUSION

In the field of transportation, one can find that the use of SP approach becomes more and more popular, especially because of its ability to investigate the responses under not-yet-existing situations (Cherchi and Ortúzar, 2006). However, in many cases, especially in practice, the SP approach has been widely applied without paying enough attention to its inherent biases caused by the hypothetical situations assumed in the SP experiment. Up to now, to correct SP biases, several types of models have been developed. This paper attempts to provide some additional models, especially focusing on behavioral dynamics and heterogeneity in the context of panel data. The behavioral dynamics could occur due to the influences of both past and future information, including initial conditions, state dependence (habit persistence and variety-seeking), and future expectation as well as time-varying tastes of trip makers. To represent the dynamics from the past information, some models like Heckman's dynamic model can be used; however, there is no model that can represent the influence of future expectation, except for

Table 6 Estimation results of Ho-DCLH model

Explanatory variables	SP Data		RP+SP Data	
	Parameter Estimates	t-stats	Parameter Estimates	t-stats
Constants term of BUS				
Parameter of 1987 (β_t)	0.2136	1.9249	0.1983	0.7721
Average annual change (Δβ)	-0.1137	-1.2057	0.0854	0.0132
Influence of time lapse on Δβ (a_t)	-0.3426	-1.0206	-1.1380	-0.0213
Parameter of 1990 (β_t)	0.1355	1.0215	0.2227	0.7850
Parameter of 1993 (β_t)	0.0740	0.5064	0.2339	1.9130
Parameter of 1994 (β_t)	0.0156	0.0972	0.2432	0.8545
Constants term of NTS				
Parameter of 1987 (β_t)	1.0922**	3.3006	8.7841	1.9113
Average annual change (Δβ)	-0.1703	-1.1233	-3.6634	-0.2637
Influence of time lapse on Δβ (a_t)	-0.1019	-1.0315	-0.1205	-0.0510
Parameter of 1990 (β_t)	0.9399**	3.2466	5.5751	1.4880
Parameter of 1993 (β_t)	0.7980**	2.6380	2.6233	1.0246
Parameter of 1994 (β_t)	0.6583	1.8547	-0.2741	-0.0848
Travel cost				
Parameter of 1987 (β_t)	-0.0017**	-4.4403	-0.0051**	-5.1885
Average annual change (Δβ)	0.0005**	2.7601	0.0005	0.1256
Influence of time lapse on Δβ (a_t)	0.1815**	3.5820	0.0900	0.0188
Parameter of 1990 (β_t)	-0.0012**	-3.7512	-0.0045**	-6.1597
Parameter of 1993 (β_t)	-0.0005	-1.5172	-0.0040**	-10.5555
Parameter of 1994 (β_t)	0.0110*	2.1116	-0.0034**	-5.2631
Travel time				
Parameter of 1987 (β_t)	-0.0144*	-2.1354	0.0002	0.0095
Average annual change (Δβ)	0.0121*	2.3247	-0.0275	-0.2314
Influence of time lapse on Δβ (a_t)	-0.2231*	-2.2034	-0.2126	-0.0836
Parameter of 1990 (β_t)	-0.0049	-1.1002	-0.0215	-1.5594
Parameter of 1993 (β_t)	0.0032	0.7830	-0.0403**	-3.6155
Parameter of 1994 (β_t)	0.0110*	2.1116	-0.0585**	-4.0426
Waiting time at bus stop or NTS station				
Parameter of 1987 (β_t)	-0.0696**	-3.1524	-0.2409	-1.3502
Average annual change (Δβ)	0.1581**	6.8576	0.3633	0.3817
Influence of time lapse on Δβ (a_t)	-0.5773**	-5.7580	-0.0689	-0.0401
Parameter of 1990 (β_t)	0.0143	0.7504	0.0960	0.4181
Parameter of 1993 (β_t)	0.0704**	3.1820	0.4171	1.9393
Parameter of 1994 (β_t)	0.1219**	4.3400	0.7348	1.8005
Access time to NTS station				
Parameter of 1987 (β_t)	-0.0424	-1.7155	-0.2764	-1.0696
Average annual change (Δβ)	0.0430	1.1341	0.1433	0.1139
Influence of time lapse on Δβ (a_t)	-0.5909**	-3.4459	-0.0628	-0.0113
Parameter of 1990 (β_t)	-0.0200	-0.9977	-0.1426	-0.4743
Parameter of 1993 (β_t)	-0.0051	-0.2020	-0.0145	-0.0686
Parameter of 1994 (β_t)	0.0085	0.2598	0.1123	0.3424
Scale parameter of SP model (μ)			0.16326*	2.27692
Initial log-likelihood	-993.15		-1619.7506	
Converged log-likelihood	-917.559		-1343.0800	
McFadden's Rho-squared	0.0761		0.1708	
Adjusted McFadden's Rho-squared	0.0637		0.1650	
Sample size	226*4=904		226*4=904	

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

Table 7 Estimation results of DCLH model

Explanatory variables	SP Data		RP+SP Data	
	Parameter Estimates	t-stats	Parameter Estimates	t-stats
Constants term of BUS				
Parameter of 1987 (β_{n0})	-0.2539	-0.7413	0.1255	1.4649
Average annual change (Δβ)	0.3889**	5.0443	0.7833	0.4903
Influence of time lapse on average annual change (π)				
Constant term	0.4293	0.5031	-4.5264	-0.5160
Gender	3.4420**	4.5440	-4.8773	-0.0115
Age	-0.1242**	-4.0260	-0.0772	-1.1827
Occupation	1.0367	1.1801	-0.8297	-0.0991
Number of household members	0.3065	1.9098	2.2829	1.9211
Constants term of NTS				
Parameter of 1987 (β_{n0})	1.7304**	3.9773	5.6722	1.3887
Average annual change (Δβ)	-0.1566	-1.9107	0.4777	0.0001
Influence of time lapse on average annual change (π)				
Constant term	-5.8349**	-5.0426	-0.0629	-0.0000
Gender	0.3893	1.0275	-0.4109	-0.0003
Age	0.0188	0.9716	-0.0771	-0.0008
Occupation	3.1886**	3.1518	-0.1977	-0.0000
Number of household members	0.7325**	4.9600	-0.4821	-0.0006
Travel cost				
Parameter of 1987 (β_{n0})	-0.0014**	-2.9261	-0.0058**	-6.0589
Average annual change (Δβ)	0.0003	1.0090	0.0017	0.2677
Influence of time lapse on average annual change (π)				
Constant term	-4.1816**	-3.8716	-3.8890	-0.0738
Gender	1.9616**	3.3656	1.8553	1.3258
Age	-0.0839**	-3.1180	-0.1645*	-1.9608
Occupation	2.0404*	2.1122	3.6545	0.0677
Number of household members	1.2074**	6.8948	1.4702*	2.2068
Travel time				
Parameter of 1987 (β_{n0})	0.0113*	2.2907	-0.0130	-1.4315
Average annual change (Δβ)	-0.0285**	-8.4576	-0.1218	-0.4257
Influence of time lapse on average annual change (π)				
Constant term	-1.1884	-1.4657	1.1480	0.2868
Gender	1.4277**	7.0709	-0.5176	-0.7882
Age	-0.0275**	-3.0376	0.0052	0.3675
Occupation	0.4488	0.5703	-0.7243	-0.3630
Number of household members	0.4863**	3.6401	-0.4329	-0.9535
Waiting time at bus stop or NTS station				
Parameter of 1987 (β_{n0})	0.0085	0.2598	-0.3263	-1.1880
Average annual change (Δβ)	0.0174	1.8813	0.3669	0.4827
Influence of time lapse on average annual change (π)				
Constant term	0.9439	1.2153	0.2343	0.0057
Gender	-0.1172	-0.2803	1.2200*	2.0989
Age	-0.0466**	-5.1743	-0.0664*	-2.3170
Occupation	1.3946	1.7866	1.5150	0.0367
Number of household members	0.0200	0.1463	0.2169	1.5319
Access time to NTS station				
Parameter of 1987 (β_{n0})	-0.0356	-1.2376	-0.0944	-0.3389
Average annual change (Δβ)	0.0015	0.1845	0.3961	0.0002
Influence of time lapse on average annual change (π)				
Constant term	-0.1101	-0.1134	0.1172	0.0000
Gender	3.8717**	4.1510	-0.5317	-0.0015
Age	-0.0877**	-3.1738	-0.0941	-0.0042
Occupation	0.7525	0.7767	-0.1527	-0.0002
Number of household members	-0.1071	-0.3886	-0.5401	-0.0031
Scale parameter of SP model (μ)			0.1234	1.8738
Initial log-likelihood	-993.1500		-1619.7506	
Converged log-likelihood	-865.4330		-1310.4200	
McFadden's Rho-squared	0.1286		0.1910	
Adjusted McFadden's Rho-squared	0.1149		0.1779	
Sample size	226*4=904		226*4=904	

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

Table 8 Estimation results of DGEV-H model

Explanatory variables	SP Data		RP+SP Data	
	Parameter Estimates	t-stats	Parameter Estimates	t-stats
Constant term of BUS				
Initial value	-0.3640	-1.6379		
Average annual change (Δβ)	0.4122**	13.6382	0.2054	0.8190
Influence of 1987 year on average annual change	0.5288	1.0351	-0.0989	-0.3416
Influence of time lapse on average annual change				
Constant term	-1.5646	-1.9047		
Gender	1.1042**	6.3932	-0.9821	-0.9863
Age	-0.0429**	-6.8423	0.0415	1.1231
Occupation	1.7787*	2.2614	-1.3377	-1.5462
Number of household members	0.2150**	3.8497	-0.0543	-0.2418
Travel cost				
Initial value	-0.0017**	-12.4753		
Average annual change (Δβ)	0.0043**	11.0663	0.1605	0.0079
Influence of 1987 year on average annual change	0.3065**	6.0970	5.1314	0.0079
Influence of time lapse on average annual change				
Constant term	0.1165	0.1434		
Gender	1.0527**	2.8986	-1.5972	-0.0798
Age	-0.0062	-0.4503	0.0060	0.0070
Occupation	-2.0381**	-3.4508	2.2006	0.0233
Number of household members	-0.4082*	-2.2626	-2.2573	-0.1139
Travel time				
Initial value	0.0019	1.1578		
Average annual change (Δβ)	-0.0038*	-2.5339	-0.0012	-1.4642
Influence of 1987 year on average annual change	0.1178	1.1034	0.2343	0.4997
Influence of time lapse on average annual change				
Constant term	-0.1075	-0.1541		
Gender	1.2476**	4.0417	0.6451	1.8827
Age	-0.0542**	-5.9102	-0.0133	-1.5216
Occupation	0.4201	0.5830	-0.1719	-0.5267
Number of household members	0.3550**	4.4524	0.1901*	2.2040
Constant term of NTS				
Initial value	0.3605	1.5914		
Average annual change (Δβ)	0.0012	0.9604	-0.0027	-0.8538
Influence of 1987 year on average annual change	1.9988*	2.1067	0.0519	0.1325
Influence of time lapse on average annual change				
Constant term	-0.8418	-0.9560		
Gender	1.4842**	2.5819	0.7282	1.2561
Age	-0.1330**	-6.6541	0.0160	0.7742
Occupation	0.5688	0.6808	-1.4546	-1.0607
Number of household members	1.9440**	8.9107	0.5040**	3.0773
Waiting time at bus stop or NTS station				
Initial value	0.2875**	5.6207		
Average annual change (Δβ)	0.5798**	19.4847	0.1503	0.4133
Influence of 1987 year on average annual change	-0.4850**	-4.9781	-0.1506	-0.4788
Influence of time lapse on average annual change				
Constant term	0.1877	0.1917		
Gender	-0.5682	-0.6263	0.8729	1.8635
Age	-0.0223	-0.5565	-0.0633*	-2.4377
Occupation	-0.7221	-0.7144	1.5481	0.9456
Number of household members	-1.9021**	-2.6077	0.3093*	2.0456
Access time to NTS station				
Initial value	-0.0097	-0.9109		
Average annual change (Δβ)	0.0017	1.2505	0.0001	0.0048
Influence of 1987 year on average annual change	0.1924	0.6045	-2.0603	-0.0022
Influence of time lapse on average annual change				
Constant term	1.0471	1.4325		
Gender	0.0271	0.1172	0.3749	0.0075
Age	-0.0078	-1.0311	0.0650	0.0449
Occupation	0.1518	0.2058	-3.1363	-0.2658
Number of household members	-0.1530**	-2.8026	0.3099	0.0210
Scale parameter of SP model (μ)			0.1261**	2.5736
Initial utility (v_0)				
BUS	0.1801	0.2694		
CAR	0.3118**	13.4353		
NTS	-0.3408	-0.5018		

Table 8 Estimation results of DGEV-H model (Continued)

Explanatory variables	SP Data		RP+SP Data	
	Parameter Estimates	t-stats	Parameter Estimates	t-stats
State dependence (ρ_{nt})				
BUS	0.3393**	18.7505		
CAR	-0.1668	-0.2523		
NTS	0.3515**	15.0927		
Influence of future expectation (φ_{nt})				
BUS				
Influence on the behavior in 1987	1.2575*	2.3035	19.3701	0.5327
Influence on the behavior in 990	5.9396*	2.4350	1.2544	1.0362
Influence on the behavior in 1993	0.9150	1.6841	0.2026	0.3151
Influence on the behavior in 1994	1.1242*	2.2417	0.0777	0.1411
CAR				
Influence on the behavior in 1987	1.1276**	3.4487	24.1484	0.5624
Influence on the behavior in 1990	0.0086	1.3722	1.3477	1.1634
Influence on the behavior in 1993	5.5361	1.7509	0.0589	0.0995
Influence on the behavior in 1994	1.3530*	2.0829	0.0404	0.0735
NTS				
Influence on the behavior in 1987	1.6527**	2.8945	0.9740	0.2413
Influence on the behavior in 1990	0.7734	1.6005	4.1223	1.2283
Influence on the behavior in 1993	0.7377	1.5091	0.1812	0.1468
Influence on the behavior in 1994	2.5036**	2.5944	2.3124	1.6110
McFadden's Rho-squared	0.3157		0.2060	
Adjusted McFadden's Rho-squared	0.2618		0.1914	
Sample size	226*4=904		226*4=904	

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

the DGEV model proposed by Swait *et al.* (2004). Future expectation is particularly important in the SP approach, because the SP experiment attempts to explore people's responses in the uncertain future situations. To reflect the behavior change over time, this paper first proposes to decompose the individual taste parameter in choice model at time point *t* into the taste at time *t-1* and change of the taste at time *t*. The taste in time *t* is re-defined as the summation of the taste at initial time and the accumulated influences of change of the taste per unit of time.

This is done by assuming that the change of the taste per unit of time is stable, but might show different influences on travel behavior due to the elapse of time. The accumulated influences are also allowed to differ across individuals. Such behavioral mechanisms are introduced into the DGEV model structure and results in the DGEV model with heterogeneity (i.e., DGEV-H model). For the purpose of comparisons, three alternative models are also built: the first model is the multinomial logit (MNL) model with time-varying taste parameters (i.e., TCP model), the second is the MNL model with the assumed behavioral dynamics in the DGEV-H model (i.e., DCLH model), and the third model is the DCLH model without heterogeneity across time (i.e., Ho-DCLH model).

To compare the above models, this paper focuses on travel mode choice behavior. The targeted new travel model is a new transit system in Hiroshima City, Japan, called Astramline, which was opened in 1994. The adopted data is a 4-wave panel data (1987, 1990, 1993 and 1994) with both RP and SP information. The above four models, i.e., TCP, DCLH, Ho-DCLH, and DGEV-H models, are estimated with respect to RP data, SP data, and SP/RP combined data, respectively. Empirical analysis results suggest that the DGEV-H model is superior to any other models in this case study. When applying the DGEV-H model, like in this case study, the SP biases could be significantly corrected even without introducing RP information. Model estimation results also confirm that the parameters of travel time and travel cost surely change over time. In summary, it can be concluded that the proposed DGEV-H model is effective to

capture heterogeneous dynamic in travel mode choice behavior.

As for future research issues, first, it is worth clarifying how to predict future travel demand reflecting the revealed heterogeneous behavior dynamics. Second, it is also important to examine how the proposed DGEV model could represent other types of travel choice behavior under different contexts considering that travel behavior is context-dependent. Finally, it is also necessary to collect some new panel data to further confirm the model performance.

REFERENCES

Ben-Akiva, M. and Morikawa, T. (1990) Estimation of switching models from revealed preferences and stated intentions, **Transportation Research Part A, Vol. 24 (6)**, 485-495.

Bradley, M. A., and A. J. Daly. (1997) Estimation of Logit Models Using Mixed Stated Preference and Revealed Preference Information. In **Understanding Travel Behaviour in an Era of Change,** Pergamon Press, Oxford, United Kingdom.

Cherchi, E. and Ortúzar, J. de D. (2006) Use of Mixed Revealed-Preference and Stated-Preference Models with Nonlinear Effects in Forecasting. **Transportation Research Record, No. 1977**, 27–34.

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

Louviere, J. J., D. A. Hensher, and J. D. Swait. (2000) **Stated Choice Methods: Analysis and Application**. Cambridge University Press, Cambridge, United Kingdom.

Morikawa, T. (1994) Correcting state dependence and serial correlation in RP/SP combined estimation method, **Transportation**, **Vol. 21 (2)**, 153-166.

Swait, J., Louviere, J. and Williams, M. (1994) A sequential approach to exploiting the combined strengths SP and RP data: application to freight shipper choice. **Transportation, Vol. 21**, 135-152.

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

Whitehead, J.C., Pattanayak, S.K., Van Houtven, G.L. and Gelso B.R. (2007) Combining revealed preference and stated preference data to estimate the nonmarket value of ecological service: An assessment of the state of the science, **Journal of Economic Surveys** (forthcoming).