A STUDY ON THE MARKET SEGMENTATION OF INTER-REGIONAL TRIPS WITH THE CONSIDERATION OF PASSENGERS' LATENT PREFERENCE FOR TRANSPORTATION MODES

Munenori SHIBATA, Railway Technical Research Institute, Japan, mshibata@rtri.or.jp

Daiki OKUDA, Railway Technical Research Institute, Japan, dokuda@rtri.or.jp

ABSTRACT

The study focuses on mode choice behavior on inter-regional trips, because inter-regional express train network is highly expected to play an important role for developing low carbon society by getting demand back from other modes which consume fossil fuels. Firstly, the paper indicates most travellers are mode captive, recognize only one transportation mode as an alternative on their mode choice behavior and latent preference factors have highly impact on the generation of mode captive. Thus, the paper tries to develop PLCS (Parameterized Logit Captivity and Selectivity) model to describe mode choice behavior more accurately, which can segment the inter-regional transportation market appropriately.

Keywords: Inter-Regional Transportation, Mode Captive, PLCS Model, Market Segmentation

1 INTRODUCTION

Inter-regional express train network is highly expected to play an important role on the inter-regional transportation system for developing low carbon society, because of its potential of low CO_2 emission and high energetic efficiency. To improve the environmental friendliness of railway, making efforts to bring the demand back to inter-regional express trains under a strict competition among other modes which consume fossil fuels, such as an airplane, an automobile, an inter-city express bus, is one of the important issues. On the other hand, as for the recent transportation policy in Japan, Japan Railway companies, which operate inter-regional high speed train (Shinkansen) and limited express trains on ordinary trucks, can supply more variety of services to passengers than before, because of the deregulation of the transportation fields by the Government. Therefore, it becomes necessary to grips the characteristics features of the market, travellers' preferences, more accurately in order to plan several services, which can adequately respond to travellers' preference.

In Japan, the current database used in most analyses and project evaluations related to Inter-regional transportation planning usually comes from the Inter-regional Net Flow Survey ("INFS" from now on) Data (e.g. Okumura et. al. (2007)). The data is based on a nationwide trip survey conducted every five year from 1990. Collected trip data and trip based OD tables are officially disclosed by Ministry of Land, Infrastructure, Transport and Tourism ("MLITT" from now on). Traditional four step method and disaggregate behavior models (e.g. disaggregate logit model) based on the INFS Data are usually applied to the demand analyses and estimation. Most of these analyses usually assume that travellers' preference, which is usually explained by some observable Level Of Service ("LOS" from now on) factors such as travel time, travel cost, frequency of transportation modes, and impedance of transfer etc, is constant in the applicable scope of these demand models. Especially, there are some assumption of applying disaggregate behavior model based on consumer behaviour theory such as disaggregate logit model to travellers' behavior. (1) Every trip maker has a perception of multiple transportation modes as choice alternative. (2) Every trip maker chooses transportation mode form multiple alternatives in order to maximize travellers' utility referring some sort of information about LOS of these alternatives.

However, it seems that there are varieties of travellers' preference segments which can be hardly explained only by abovementioned LOS factors, because an inter-regional trip, which is different from inner-regional daily trips such as commuter trips etc, is a rare experience for most trip makers and information amount of inter-regional trips seems to be smaller than that of daily trips. In other words, there are some possibilities that the decision making is influenced not only by the LOS, but also by magnitude of importance for intangible factors, such as punctuality, easiness of loading baggage etc (e.g. Shibata et. al. (2001)). Additionally, travellers' potential preference for inter-regional transportation modes, such as like or dislike for each transportation mode itself, seems to effect to the mode choice behavior when a trip purpose is non-business.

Thus, the study focuses on the varieties of the abovementioned travellers' intuitive subjective factors, which are named "latent preference factors" from now on in the study, and the influences of them to the mode choice behavior on Inter-regional travellers. A pilot survey on an experience of mode choice behavior in an inter-regional transportation market is conducted. The trip survey is composed of the questions not only about some trip attributes, but also about latent preference factors, such as categorized magnitude of travellers' potential preferences for inter-regional transportation modes, and magnitude of importance for qualitative characteristics of each mode. In the paper, some analyses indicate characteristics of the relationship between mode choice result, choice alternatives and magnitude of travellers' latent preference factors, which suggests that the inter-regional transportation market is divided into some segments by categories of choice alternatives. Finally, the study also tries to calibrate the PLCS (Parameterized Logit Captivity and Selectivity) model to describe the market segmentation demonstratively on the individual mode choice model. The purpose of the study is to demonstrate the importance of applying the market segmentation based on travellers' latent preference to analyze the mode choice behavior of inter-regional trips more accurately.

2 SURVEY CONDUCTION OF INTER-REGIONAL TRIPS

2.1 Findings from Previous Studies

In the past decade, some studies to analyse and develop appropriate models of mode choice behavior of inter-regional trips have been made in Japan. For example, Muto, M. et. al. reveal that mode choice behavior between an inter-regional express train and an automobile on inter-regional trip makers with non-business purpose are strongly influenced by whether a traveller attaches importance to some qualitative characteristics of each mode or not. For example, trip makers who put priority on punctuality and safety strongly tend to use express trains and who put priority on easiness of loading baggage and mobility strongly tend to use an automobile (Muto et. al. (1999)). After that, Muto, M. et. al. suggest that after mentioned modelling methodology developed by Morikawa et. al. (1993) is efficient to describe the mode choice behavior (Shibata et. al. (2001), Muto et. al. (2004)). (1) Collecting the magnitude of importance for intangible factors, such as qualitative characteristics of each mode, from trip makers in trip survey. (2) Modelling the relationship between these magnitude of importance for intangible factors and objective observable data (e.g. trip attributes) by the Multiple Indicator Multiple Cause (MIMIC) model, which is a kind of Structural Equation Modelling (SEM). (3) Fitted latent values estimated by the MIMIC model are introduced into utility functions of disaggregate mode choice models.

Recently, the study focused on choice alternatives of mode choice behavior on inter-regional travellers has been attempted, because of the high impact of magnitude of importance for intangible factors to mode choice behavior suggested by abovementioned previous studies. Shibata et. al. (2009) have already indicated that (1) most of inter-regional travellers with non-business purpose is mode captive, who recognize only the mode actual chosen in the trip as an alternative from the collected trip survey data, in the aspect of mode choice between an inter-regional express train and an automobile, (2) Decision making process of mode choice behavior contains at least 2 steps, the alternative screening process and mode choice determining process. (3) The screening process is highly influenced by latent values based on some magnitude of importance for intangible factors. (4) The Parameterized Logit Captivity (PLC) model developed by Swait et. al. (1987) can described these processes accurately, because the model can abstract mode captive segments from the travellers' market. Then, M-PLC model, the PLC model with the MIMIC model, is proposed.

2.2 Outline of the Trip Survey

Therefore, the study conducts the original questionnaire survey in order to collect trip data of every representative inter-regional transportation mode, inter-regional express train passengers, inter-city express bus passengers, airplane passengers, and automobile users in November 2008. The survey aims to collect not only trip attributes data (e.g. origin, destination, the result of choice (=revealed preference) etc), but also mode alternatives recognized by a trip maker, and some magnitude of importance for intangible factors. Additionally, the study also focuses on travellers' potential preference for inter-regional transportation modes, such as like or dislike for each transportation mode itself, because these seem to effect to the mode choice behavior when a trip purpose is non-business. So, magnitude of potential preference for each mode is also observed from each trip maker independent from the section of trip survey.

The trip survey is conducted through the web survey system, and the respondents are the driver's licence holders who made domestic inter-regional travel within past two months. In Japan, prefecture is local government administrative area. There are 47 prefectures in Japan. Inter-regional travel is basically defined that trips over the prefectural boundary. In case of trips originate out of three major metropolitan areas (Greater Tokyo including Tokyo metropolitan, Kanagawa prefecture, Saitama prefecture and Chiba prefecture, Greater Nagoya including Aichi prefecture, Gifu prefecture and Mie prefecture, Greater Osaka including Osaka prefecture, Kyoto prefecture, Hyogo prefecture and Nara prefecture.), trips over the greater metropolis area boundary is objective for the trip survey. These definitions are same as INFS.

Additionally, residents in Hokkaido prefecture and Okinawa prefecture, which are not connected by road network to Honshu (main land), Kyushu and Shikoku, are excluded in the candidates of respondents of the survey, because these residents hardly include an automobile in their mode alternatives set. Similarly, trips whose destination is in Okinawa prefecture (South western isolated islands) are excluded, because most of travellers have no choice but to use airplanes to these isolated islands. Residential area of the respondents is shown in Figure 1.

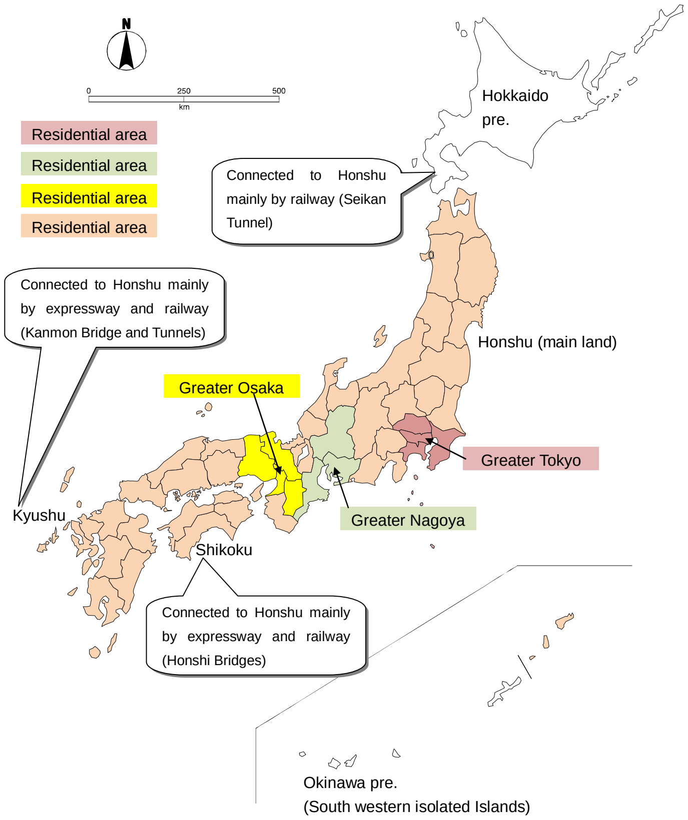

Figure 1 – Residential Compound of the Respondents

Table 1 – Main Contents of the Questionnaire of the Trip Survey

potential preference for each mode	five ranked (+1~+5) like or dislike of inter-city express train, inter-city express bus, airplane and automobile
trip experience & revealed preference	trip purpose, date of departure, itinerary origin and destination travelling route, access modes, line haul mode, egress modes the number of accompanies and its category, fare
mode alternative	Inter-regional transportation mode recognized as alternatives by the respondent (inter-city express train, inter-city express bus, airplane, automobile, others)
magnitude of importance for intangible factors	seven ranked (+1~+7) magnitude of importance for intangible factors as below (21 items) ability of reaching rapidly, low travel cost, punctuality select-ability of departure time, familiarity of chosen transportation mode harmony with the environment (ex. low CO_2 emission),mobility easiness of loading baggage, possibility of meeting traffic congestion on road safety for traffic accident, safety for security, fatigue by driving automobile reserving private space, effective utilization of travelling time impedance of transfer, accessibility for railway station/bus stop/airport accompanied with children/elderly person, enjoy-ability of driving automobile enjoy-ability of boarding trains, enjoy-ability of boarding airplanes enjoy-ability of boarding inter-city express buses
individual attributes	gender, age, occupation, car ownership, possession of Electric Toll Collection (ETC) instrument admission of transport enterprises' membership (airline mileage membership etc.)

The main contents of the questionnaire of the trip survey are shown in Table 1. The questionnaire mainly aims to observe these three contents. (1) Observing five ranked potential preference (+1~+5) such as like or dislike of each mode independently from the after-mentioned contents as for trip attributes. (2) Observing not only the result of mode choice, but also alternative modes recognized by a trip maker. (3) Observing magnitude of seven ranked (+1~+7) importance for intangible factors (21 items) in an aspect of mode choice behavior.

2.3 Outline of the Data Set for Analyses

As mentioned previously, Shibata et. al. (2009) has already proposed model named M-PLC model to describe the decision process of mode choice behavior. However, M-PLC model is for the mode choice behavior between an inter-regional rapid train and an automobile (only 2 alternatives). On the other hand, inter-city express bus (Coaches) is one of the strong competitors from the practical viewpoints of railway companies in some regions (e.g. Shibata et. al (2006)). Therefore, the study tries to analyses the decision making process of mode choice behavior between an inter-regional express train, an automobile and an inter-city express bus (3 alternatives). In other words, the study aims to expand the applicable scope of previously proposed M-PLC model.

Sample data is extracted according to the criteria shown as below, (1) Trips of inter-regional express train users, automobile users and inter-city express bus users is object of the study. (2) Trips in which an airplane is recognized as an alternative by the trip maker are excluded. (3) Trips in which an inter-city express bus cannot be used because of the nonexistent of the express bus line are excluded. Above criteria brings the final sample size to 1,611. Free address matching service on the web operated by the Center for Spatial Information Science, the University of Tokyo and National Integrated Transport Analysis System (NITAS) developed by the MLITT (2009) is applied to create some portion of LOS data (e.g. travel time, cost, frequency etc.) for each sample.

Figure 2 shows the data profile of the objective trip data set. The survey can observe every category of travellers almost evenly, thus, the data set seems to be valid for the study. Figure 3 indicates the relationship between one way trip distance and the share of modal split in the data set. Basically, the share of express train increases with increasing the trip distance, with correspondent to the mode share in Japan observed by the INFS. On the other hand, the share of express bus is decreasing constantly from about 20% (category; ~299km) to about 15% (category; 500~699km). Vice versa, the share of express bus is increasing in case of trip distance is over 700km. In recent years, most of overnight express trains with sleeper are abolished. On the other hand overnight express bus network is expanding. Above circumstances of overnight transport services seems to be reflected in the observed sharing tendency.

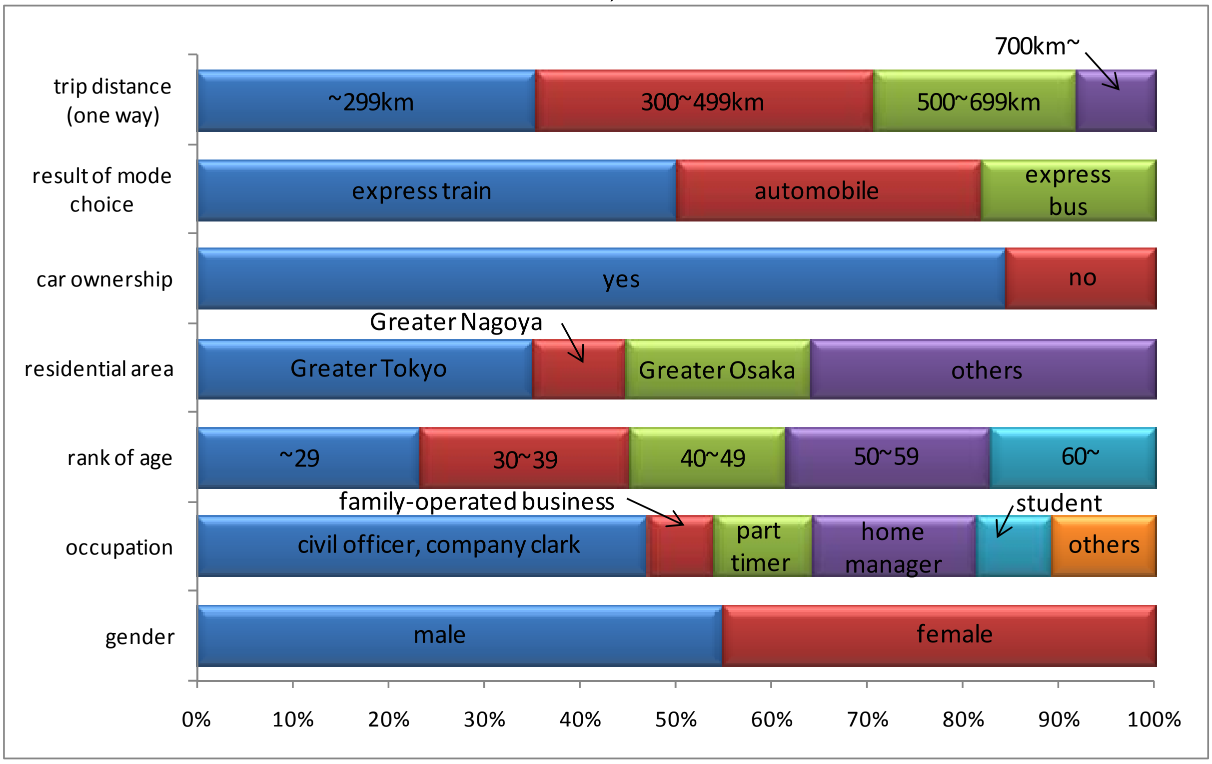

Figure 2 – Data Profile for Analyses and Modeling (N=1,611)

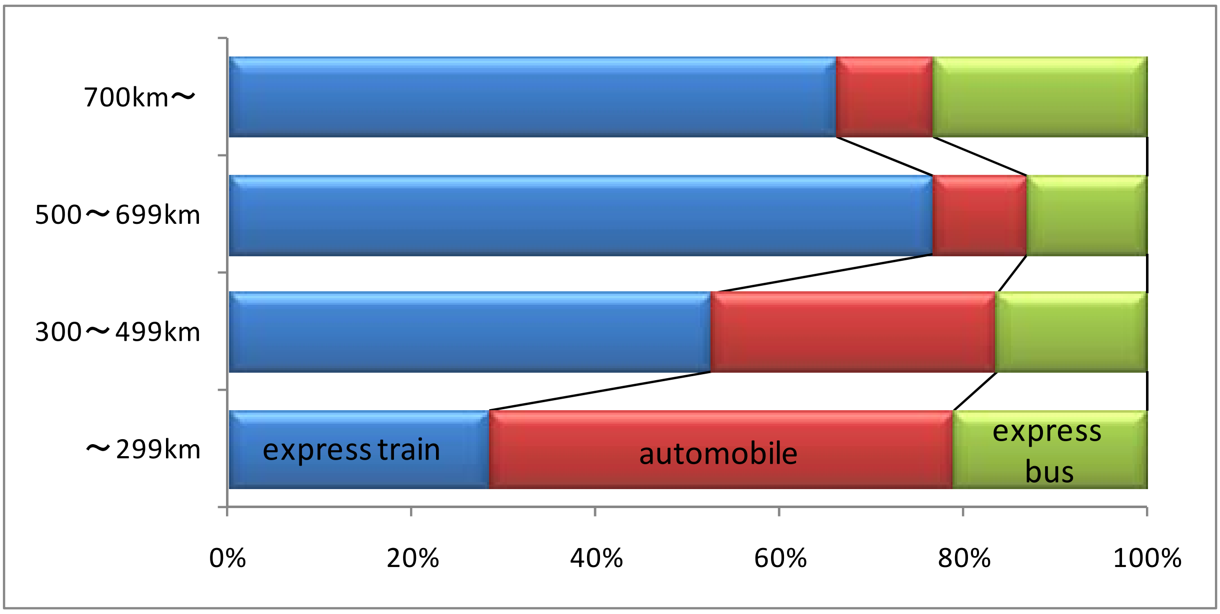

Figure 3 – Relationship between One Way Trip Distance and Share of Modal Split of the Data Set (N=1,611)

3 INITIAL FINDINGS

Initially, the relationship between the mode choice behavior and the observed latent preference should be confirmed. In this chapter, some aggregate and disaggregate analyses are conducted in order to reveal the basic characteristics of the mode choice behavior among 3 modes. In these following analyses in this chapter, seven ranked categorized data of magnitude of importance for intangible factors and five ranked categorized data of potential preference for each mode are normalized by the successive category method (Ikeda (1986)) often applied in the fields of psychology, for example categorical data processing of sensory inspection. The converted dataset by the successive category method can be qualified as metric variables according with standard normal distribution (average: 0, standard deviation: 1).

3.1 Effectiveness of Latent Preference to Mode Choice

Figure 4 indicates that mean value of relative potential preference of each transportation mode. The relative potential preference is defined as a difference between the potential preference for one mode and the average of all potential preferences, which are calculated per sample. The mean values of relative potential preference are averaged by each transportation mode user. Each mode users has high potential preference for chosen transportation mode relatively. Figure 5 shows the some discriminative intangible factors' mean value of relative magnitude of importance for 21 intangible factors. Same as above, the relative magnitude of importance for the 21 mode characteristics is defined as a difference between the intangible factor and the average of all intangible factors, which are calculated per sample. The mean values of relative magnitude of importance are averaged by each transportation mode user. As for mode qualitative characteristics, express train users place weight particularly on "ability of reaching rapidly" and "punctuality". On the contrary, automobile users give more importance to "mobility", "easiness of loading baggage" and "reserving private space". Inter-city express bus users enormously put high priority on "low travel cost". Regarding enjoy-ability, each mode users have high weight relatively on enjoy-ability for chosen transportation mode, for example, train users have high magnitude of importance on "enjoy-ability of boarding trains". This suggests that the view point of treating transportation mode as one of hobbies, such as "Rail fan", may effect to the mode choice behavior to some extent.

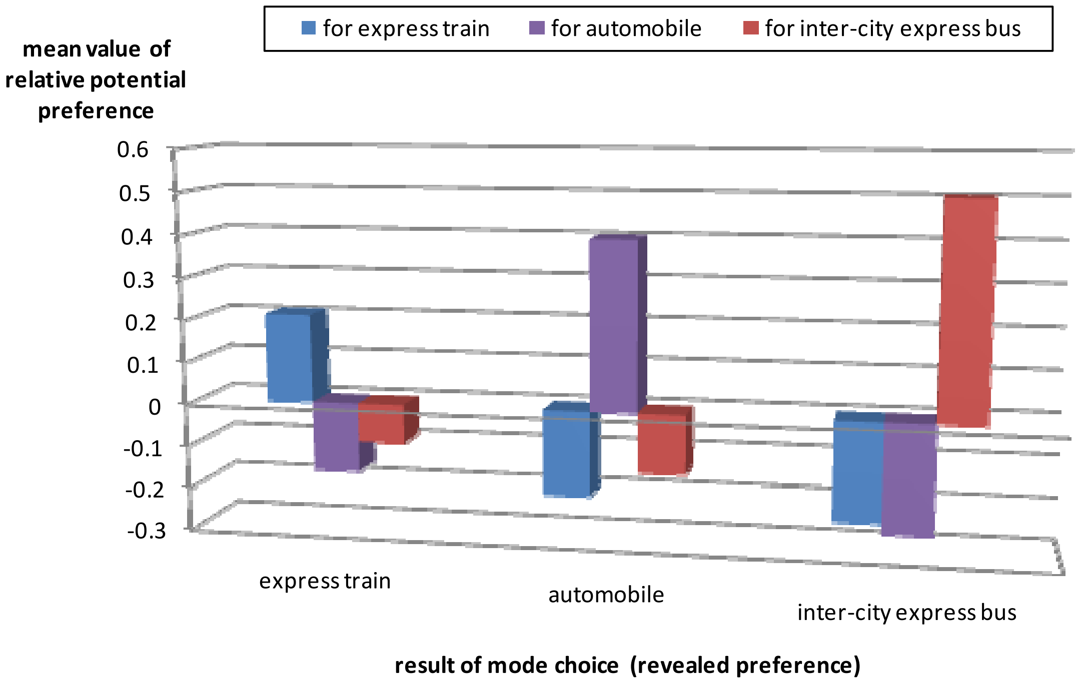

Figure 4 – Mean Value of Relative Potential Preference for Each Transportation Mode by Result of Mode Choice

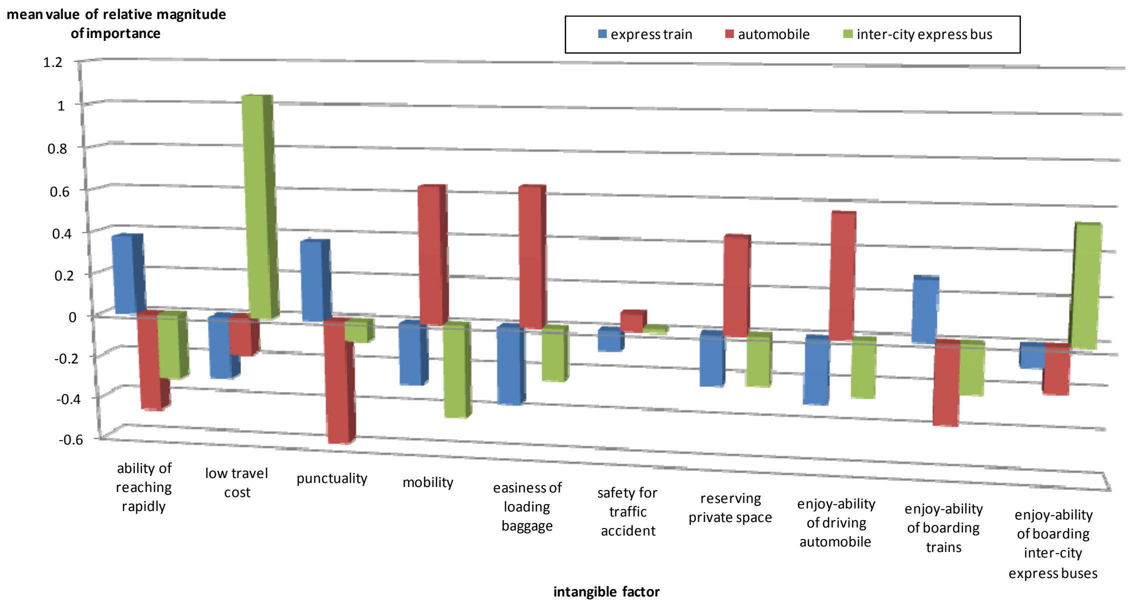

Figure 5 – Mean Value of Relative Magnitude of Importance for Intangible Factors by Result of Mode Choice

In order to grips above characteristics statistically, some mode choice models are calibrated based on the traditional disaggregate logit model formulated by equation (1) and (2).

$$P_{in} = \frac{\exp(V_{in})}{\sum_j \exp(V_{jn})} \qquad (1)$$

$$V_{in} = \sum_k \theta_{ink} X_{ink} \qquad (2)$$

where, P_{in} : probability that individual n selects transportation mode i , V_{in} : utility function of mode i , X_{ink} : each explanatory variable, θ_{ink} : parameters to be estimated. The estimated parameters of logit models for mode choice behavior are shown in Table 2. The model 1 has only common variables of LOS and mode constant in utility functions. On the other hand, normalized magnitude of importance for various intangible factors and normalized potential preference for each transportation mode are installed as mode specific variables on utility functions in the model 2.

Each parameter of both models except for a part of mode constant has significant statistical absolute t-value, larger than 1.96. As for the sign condition, parameters of "travel time" and "travel cost" should be negative as usual, however, each parameter of normalized latent preference variable seems to be positive according to the above mentioned aggregate analyses (Figure 4, Figure 5). For example, travellers who put higher importance on "punctuality" are expected to tend to choose inter-city trains. Each signs of these parameters is corresponding to the each sign condition. As for the goodness of fit of these models shown by $\bar{\rho}^2$, both models have significant $\bar{\rho}^2$, larger than 0.2. However, model 2 is better by far than the model 1 due to introducing latent preference variables. These indicate statistically that mode choice behavior is influenced not only by "travel time" and "travel cost" (LOS), but also by latent preference variables also in the dataset of the study.

Table 2 – Estimated Parameters of Logit Models for Mode Choice Behavior

variables			model 1	model 2
common variable		travel time (hour)	-0.5092 (-20.03)	-0.5374 (-11.47)
		travel cost (10,000yen/person)	-2.0045 (-12.98)	-1.4836 (-6.456)
magnitude of importance for intangible factor (normalized)	R	ability of reaching rapidly		1.0590 (9.411)
	R	punctuality		0.9406 (8.703)
	A	mobility		1.0960 (7.042)
	A	easiness of loading baggage		1.5095 (9.107)
	A	reserving private space		0.6889 (5.348)
	B	low travel cost		2.1568 (15.42)
potential preference (normalized)	R	for express train		0.6449 (6.523)
	A	for automobile		0.7983 (6.923)
	B	for inter-city express bus		0.9911 (8.430)
mode constant	R	express train constant term	0.0174 (0.306)	0.5627 (2.474)
	A	automobile constant term	-0.6156 (-6.917)	-0.4651 (-2.511)
$\bar{\rho}^2$			0.215	0.643
AIC (Akaike's Information Criteria)			2784.09	1271.69
hit ratio of mode choice result (%)			63.6%	85.1%
value of time (yen/minute)			42.3	60.4
number of samples			1,611	1,611

* ( ) : statistical t-value
* R: express train specific variable A: automobile specific variable B: inter-city express bus specific variable

3.2 Effectiveness of Latent Preference to Alternative Screening

Observing the mode alternative recognized by the trip maker is one of the important features of the trip survey of the study. On the basis of abovementioned collected mode alternative data, relationship between mode alternative, LOS and latent preference are analyzed in this section. In the followed analyses, a mode captive is defined that a traveller recognizes only the chosen transportation mode as an alternative. On the contrary, for the sake of simplicity of the followed analyses and modelling, a selective is defined that a traveller recognizes more than two modes including actual chosen mode. Then, there are four categories of mode alternative in the study, mode captive for an express train (Rcap), mode captive for an automobile (Acap), mode captive for an inter-city express bus (Bcap), and selective (sel).

Figure 6 shows the share of mode alternative categories aggregated by each mode user. The share of mode captive usually keeps high sharing rate around 90 % in every mode user. This seems to suggest that analysing characteristics of screening process of mode

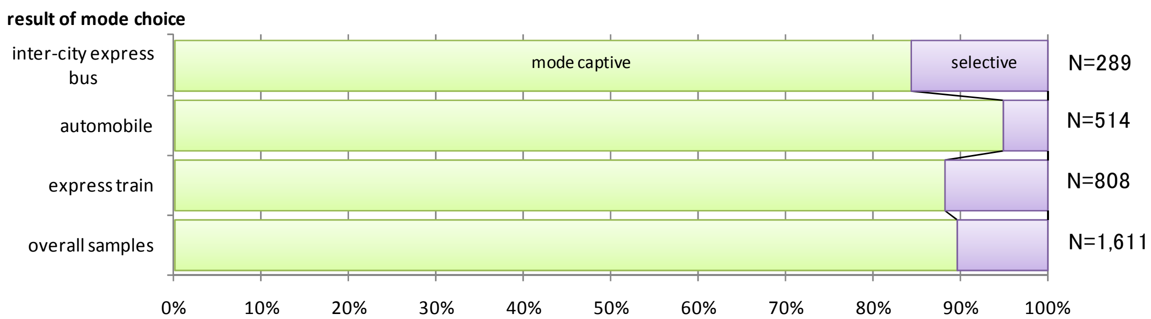

Figure 6 – Observed Share of Mode Alternatives Categories

alternative should be done as a preparation for modelling of mode choice behavior.

Figure 7 indicates that mean value of relative potential preference of each alternative category. In the same way of **3.1**, the relative potential preference is defined as a difference between the potential preference for one mode and the average of all potential preferences, which are calculated per sample. The mean values of relative potential preference are averaged by each alternative category. Each captive category has high potential preference for its transportation mode relatively. On the other hand, potential preferences of selective category are average point around 0 in comparison with captive categories. Figure 8 shows the some discriminative intangible factors' mean value of relative magnitude of importance for 21 intangible factors same as **3.1**. The relative magnitude of importance for the 21 mode characteristics is defined as a difference between the intangible factor and the average of all intangible factors, which are calculated per sample. The mean values of relative magnitude of importance are averaged by each alternative category. Regarding mode qualitative characteristics, Rcap places weight particularly on "ability of reaching rapidly" and "punctuality". Acap gives more importance to "mobility", "easiness of loading baggage" and "reserving private space". Bcap enormously puts high priority on "low travel cost". Regarding enjoy-ability, each captive category has high weight relatively on enjoy-ability for its transportation mode, for example, Acap has high magnitude of importance on "enjoy-ability of driving automobile". On the contrary, selective category has average magnitude of importance around 0 on most intangible factors in comparison with captive categories.

Above-mentioned characteristics of aggregate analyses suggest that neutral travellers on both potential preference and magnitude of importance for intangible factors tend to belong to selective category, and travellers who have high importance on potential preference on specific mode tend to belong to specific mode captive. Similarly, travellers with high magnitude of importance on intangible factors of specific mode characteristics increase tendency of belonging to the relevant mode captive.

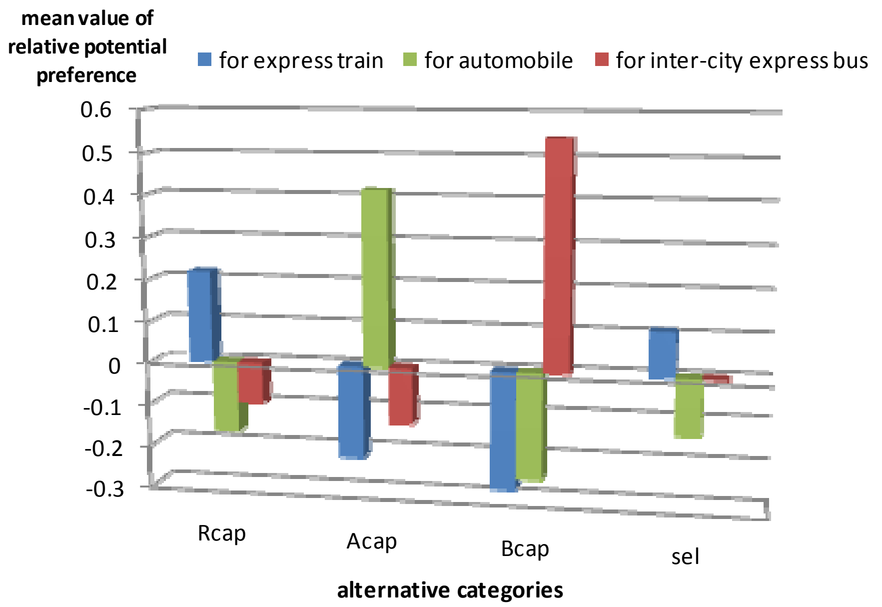

Figure 7 – Mean Value of Relative Potential Preference for Each Transportation Mode by Each Alternative Category

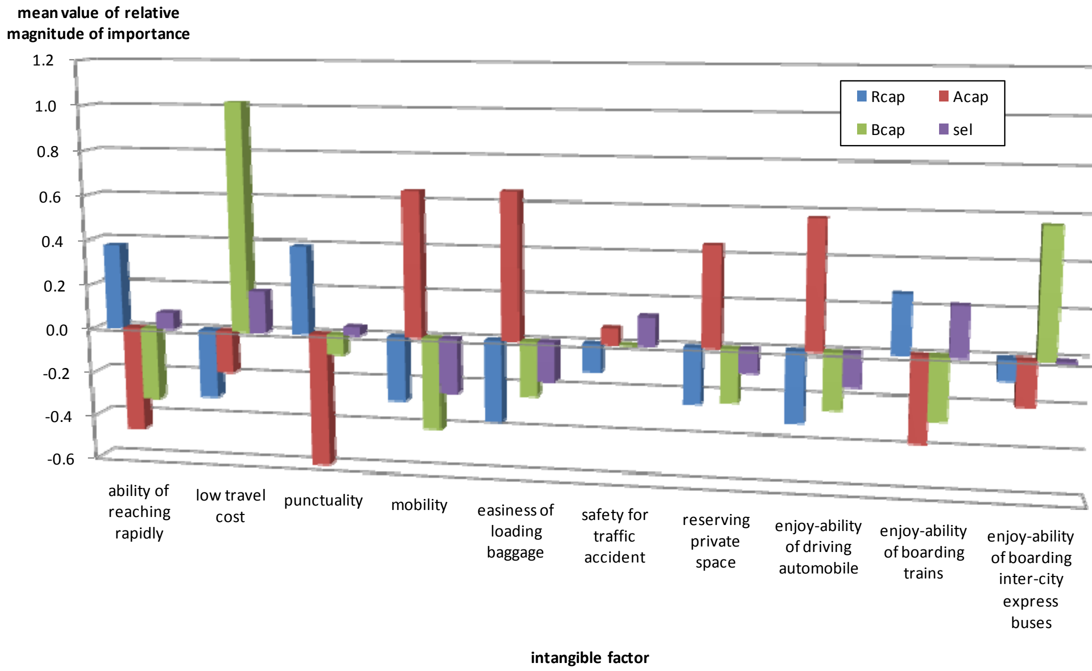

Figure 8 – Mean Value of Relative Magnitude of Importance for Intangible Factors by Each Alternative Category

In order to verify characteristics statistically found by above aggregate analyses statistically, some alternative category probability models are calibrated based on the traditional disaggregate logit model formulated by equation (3) and (4).

$$P_{in} = \frac{\exp(V_{in})}{\sum_j \exp(V_{jn})} \qquad (3)$$

$$V_{in} = \sum_k \theta_{ink} X_{ink} \qquad (4)$$

where, P_{in} : probability that individual n belongs to alternative category i , V_{in} : function which explains belonging to alternative category i , X_{ink} : each explanatory variable, θ_{ink} : parameters to be estimated. However, function of selective category is fixed as 0 in order to estimate parameters of other functions. The estimated parameters of alternative category probability models are shown in Table 3.

In the Table 3, GC means Generalized Cost which is calculated by the summation of travel cost and cost of travel time corresponded by the value of time of mode choice model 1 in the table 2. Min. difference of GC is defined that the difference of relevant mode's GC and GC of most competitive mode. Here, most competitive mode means the mode which has nearest GC to the relevant mode. For concrete example, min. difference of RGC is calculated by GC of express train minus GC of automobile when GC of automobile is nearer than GC of inter-city express bus. If RGC is larger than most competitive mode's GC, probability of Rcap is expected to decrease. Therefore, the sign condition of each minimum (min.) difference of GC should be negative.

In the model 1, category probability is explained only by min. difference of GC, in other words, only by difference of LOS. The model 2 expresses the category probability only by the

Table 3 – Estimated Parameters of Alternative Category Probability Models

variables			model 1	model 2	model 3
min. difference of GC	R	min. difference of RGC	-0.6688 (-5.169)		-0.5882 (-3.937)
	A	min. difference of AGC	-3.0364 (-15.14)		-2.6794 (-10.46)
	B	min. difference of BGC	-0.8731 (-5.330)		-0.7432 (-4.107)
magnitude of importance for intangible factor (normalized)	R	ability of reaching rapidly		0.4608 (5.837)	0.5238 (6.365)
		punctuality		0.6424 (8.152)	0.5270 (6.478)
	A	mobility		0.8830 (7.282)	0.8549 (6.336)
		easiness of loading baggage		1.1404 (8.765)	1.0521 (7.488)
		reserving private space		0.5010 (5.029)	0.4812 (4.353)
	B	low travel cost		1.4442 (12.82)	1.4243 (13.21)
potential preference (normalized)	R	for express train		0.3868 (5.427)	0.3444 (4.632)
	A	for automobile		0.6980 (7.582)	0.6325 (6.090)
	B	for inter-city express bus		0.8251 (8.229)	0.8001 (8.140)
constant term	R	Rcap constant	1.3357 (14.91)	1.4794 (16.20)	1.3499 (14.88)
	A	Acap constant	0.3138 (2.706)	0.5501 (4.827)	-0.1583 (-1.089)
	B	Bcap constant	0.6923 (6.086)	-0.2343 (-1.727)	0.0245 (0.269)
$\bar{\rho}^2$			0.213	0.419	0.462
AIC (Akaike's Information Criteria)			3525.27	2604.33	2415.23
hit ratio of alternative category (%)			57.7%	72.4%	74.2%
number of samples			1,611	1,611	1,611

* ( ) : statistical t-value
* R: Rcap (express train captive) A: Acap (automobile captive) B: Bcap (inter-city express bus captive)
* Function of selective category is fixed as 0 to estimate parameters.

normalized latent preference factors. In the model 3, both difference of LOS and normalized latent preference factors are installed into the functions which explain category probability. As for goodness of fit shown by $\bar{\rho}^2$, each model has significant $\bar{\rho}^2$, larger than 0.2. However, model 2 and model 3 are better by far than model 1 due to introducing latent preference variables. Moreover, the model 3 is more significant than the model 2, because the lowest AIC (Akaike's Information Criteria) model is model 3. From the point of parameters' significance of explanatory variables, both differences of LOS and latent preference factors have significant t-value. That is to say, both differences of LOS and latent preference factors have possibility of expressing which category the traveller belongs to in some behavior models.

4 MARKET SEGMENTATION REVEALED BY THE PLCS MODEL

4.1 Assumption of Decision Making Process of Mode Choice Behavior

According to abovementioned analyses, the principal characteristics of the mode choice behavior on inter-regional travellers with non-business purpose can be summarized as follows;

A) Most travellers recognize only the transportation mode which is chosen actually as a single mode alternative. In other words, most travellers are mode captive and the share of selective is extremely small according to the collected mode alternative recognition data.
B) LOS factors have effect on both mode choice behavior and mode alternative screening with statistical efficiency.

C) However, latent preference factors such as magnitude of importance for intangible factors and potential preference for each mode have a substantial influence on both mode choice behavior and mode alternative screening.

Hence, it is necessary to treat the mode choice behavior as at least 2-step decision making process referring to Manski (1977).

(1st step) : "Mode alternative screening process", where travellers recognize single transportation mode or several other modes as choice alternatives. The former is mode captive and the latter is selective in the study.

(2nd step) : "Determining process" in which selective travellers choose only one mode among some alternatives.

In Japan, a lot of studies as for choice behavior on inter-regional travellers are already done. Most of them apply disaggregate discrete choice model to describe behavior. Namely, most studies focus only on the above 2nd step "Determining process". Very few attempts of observing, analyzing and modelling above 1st step "Determining process" have been made in the field of inter-regional travellers' behavior. The study is ranked as earlier basic stage of developing a mode choice behavior model of inter-regional traveller with respond to the actual condition of recognition of mode alternatives.

Referring to the result of analyses in chapter 3, the study tries to explain the "Screening process" by mainly latent preference factors and secondarily difference of LOS (Generalized Cost). On the other hand, it is tried to describe the "Determining process" mainly by LOS, because selective travellers are assumed that they choose the transportation mode among several alternatives more reasonably by comparing LOS of each mode.

The assumption of decision making process of mode choice behavior of inter-regional travellers with non-business purpose in the study is shown in Figure 9.

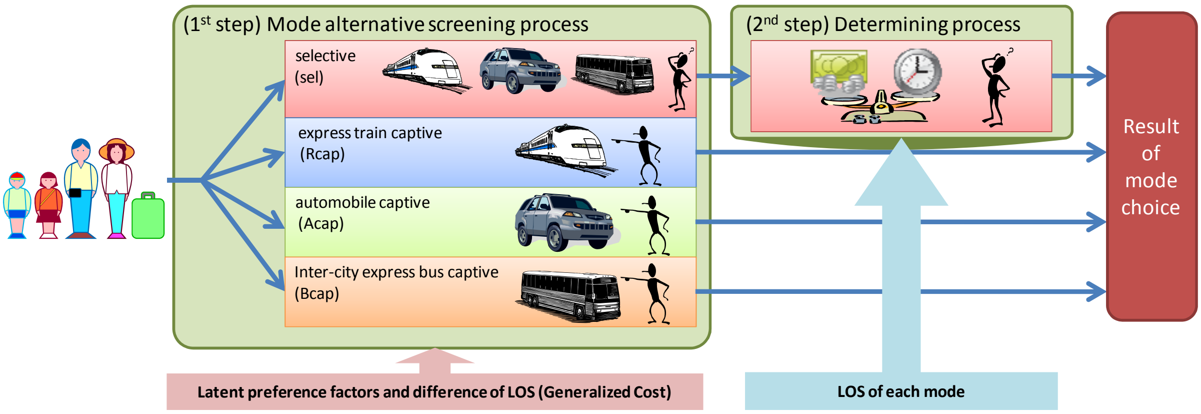

Figure 9 – Assumption of Decision Making Process of Mode Choice Behavior in the Study

4.2 The PLCS Model

The general formula of random utility model with the consideration of alternative screening process is formulated as equation (5) (Manski (1977)).

$$P_{in} = \sum_{C \in G_j} P_n(i \mid C) \cdot Q_n(C \mid G_j) \tag{5}$$

where, P_{in}: probability that individual n selects mode i, $P_n(i \mid C)$: probability that mode i is included in choice alternative of individual n, G_j: alternative set j, $Q_n(C \mid G_j)$: probability that individual n has alternative set C screening among all alternative set G. In the study, 1st step "Mode alternative screening process" corresponds to $Q_n(C \mid G_j)$ and 2nd step "Determining process" corresponds to $P_n(i \mid C)$.

4.2.1 Formulation of PLCS Model

The study deals mode choice behavior among 3 alternatives (i=1: express train, 2: automobile, 3: inter-city express bus). For the sake of simplicity of the followed modelling, a selective is defined that a traveller recognizes more than two modes including actual chosen mode in the study. Therefore, there are four categories of mode alternative. Suffix j is mode alternative category number. G is defined as $G_0 = \{(1,2,3)\}$: "sel", $G_1 = \{(1)\}$: "Rcap", $G_2 = \{(2)\}$: "Acap", $G_3 = \{(3)\}$: "Bcap".

As for the 1st step "Alternative screening process", the alternative category probability function is formulated as commonly used logit probability model in the study.

$$Q_n(G_j) = \frac{\exp(U_{jn})}{\sum_j \exp(U_{jn})} \qquad (6)$$

$$U_{jn} = \sum_k \theta_{jnk} X_{jnk} \qquad (7)$$

where, $Q_n(G_j)$: probability that individual n has alternative set G_j screening among all alternative set G, U_{jn}: function of alternative category j of individual n, X_{jnk}: each explanatory variable, θ_{jnk}: parameters to be estimated.

Moreover, in respect of 2nd step "Determining process", only when the traveller belongs to the G_0 category (selective category), whoever selects one transportation mode among several mode alternatives. Traditional disaggregate logit model can be applied because this process can be interpreted as usual choice behavior with several alternatives. 2nd step is formulated as equations (8) and (9).

$$P_n(i \mid G_0) = \frac{\exp(V_{in})}{\sum_i \exp(V_{in})} \qquad (8)$$

$$V_{in} = \sum_k \beta_{ink} Y_{ink} \qquad (9)$$

where, $P_n(i \mid G_0)$: probability that individual n selects transportation mode i when the n belongs to G_0, V_{in} : utility function of mode n, Y_{ink} : each explanatory variable, β_{ink} : parameters to be estimated.

The PLC (Parameterized Logit Captivity) model developed by Swait et. al. (1987) can expresses the belongingness to only each mode captive category by functions composed of some attribute variables in the 1st step. On the contrary, the model formulated in the study can consider the factors that lead travellers to belong to not only captive categories but also to the selective category in the 1st step. Hence, the decision making process model is named PLCS (Parameterized Logit Captivity and Selectivity) model by the authors. The model can be ranked as one of the enhancement type of PLC model.

4.2.2 Abstraction of "Attitude of Importance for Each Mode" by Factor Analysis Model

Many kinds of latent preference factors influence the alternative screening process of mode choice behavior as shown in Table 3. However, it is expected to be difficult to deal these many latent factors in a decision making process model. Hence, the study tries to epitomize these many factors into some abstracted factors. It is expected to be useful in order to install these latent factors into the PLCS model.

Factor analysis is applied to latent preference factors' data in order to abstract some epitomized factors. The result of factor analysis is shown in Figure 10. Each latent preference data is not normalized and is inputted "as is" in each observed latent factor shown by a square box in Figure 10. GFI is goodness-of-fit index of the estimated model, and AGFI is adjusted GFI. From the fitness shown by GFI and AGFI, the model can abstract some reasonable latent factors from observed latent preference factors' data with enough accuracy.

Three latent factors "Attitude of Importance (A.I. from now on) for each mode" are abstracted by the model. The latent factor #1 should be named "A.I. for express train", because it is composed with "ability of reaching rapidly", "punctuality" and "potential preference for express train". The latent factor #2 consists of "mobility", "easiness of loading baggage", "reserving private space" and "potential preference for automobile". Therefore, it is named "A.I. for automobile". Similarly, the latent factor #3 can be named as "A.I. for inter-city express bus", because it is influenced by "low travel cost" and "potential preference for inter-city express bus".

4.2.3 Result of Parameter Estimation of the PLCS Model

The PLCS model has one important feature in comparison with the PLC model. The feature is that the element of explaining the condition, where a traveller tends to belong to the selective category, is installed into the function U_{0n} of "Alternative screening process" (equation (7)). Regarding to abovementioned characteristics mentioned by the result of aggregate analyses in section **3.2**, neutral travellers on both potential preference and magnitude of importance for intangible factors tend to belong to the selective category. Thus, a special index "index of difference between each $A.I.$ ($dif.A.I.$ from now on)" is installed into the function in the study. The index $dif.A.I.$ is defined as equation (10).

$$dif.A.I._n = \frac{1}{|AI_{1n} - AI_{2n}| + |AI_{1n} - AI_{3n}| + |AI_{2n} - AI_{3n}|} \qquad (10)$$

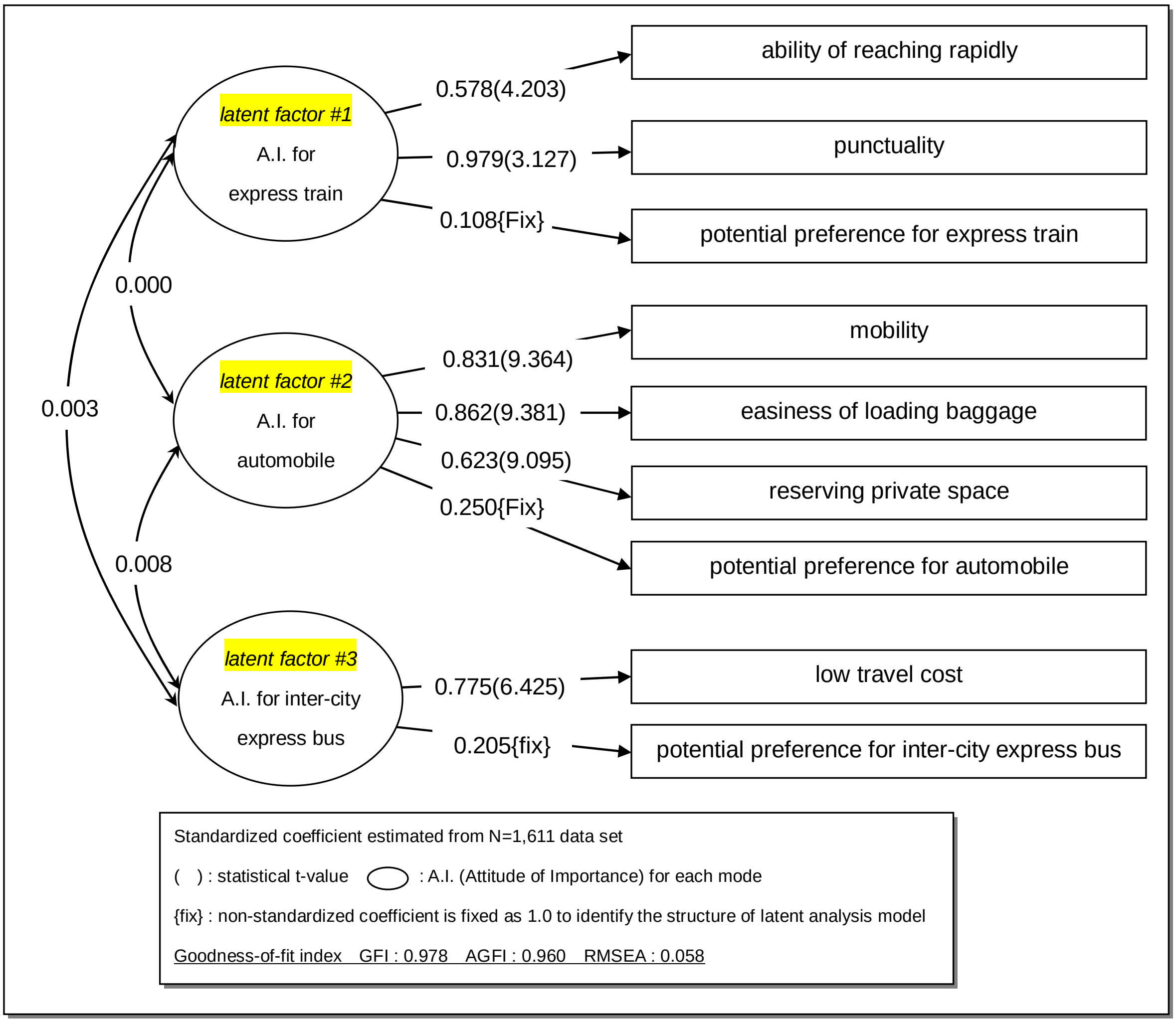

Figure 10 – Estimated Parameters of Latent Analysis Model on Latent Preference Factors

where, $dif.A.I._n$: $dif.A.I.$ index of individual n, AI_{in}: latent factor score ($A.I.$) of mode i of individual n. When every $A.I.$ has same factor score, $dif.A.I.$ reach an infinite value. On the other hand, the larger the difference of $A.I.$ becomes, the smaller $dif.A.I.$ to be calculated. The sign condition of $dif.A.I.$ should be positive in the function of the selective category, because it is assumable that probability of belongingness to the selective category seems to be larger when $dif.A.I.$ of a traveller is smaller.

The calibrated parameters of the PLCS models are shown in Table 4. In the $Q_n(G_i)$ of the model 1, AI_n, $dif.A.I.$ is installed into the functions. In the model2, "Min. difference of GC" defined in chapter **3.2** is additionally installed. The utility functions of $P_n(i\,|\,G_0)$ of both models have explanatory variables of LOS and mode constant. All parameters in both steps are estimated simultaneously.

As for goodness of fit shown by $\bar{\rho}^2$, each model has significant $\bar{\rho}^2$, larger than 0.2. However, model 2 is superior to the model 1 due to introducing min. difference of GC. Moreover, the AIC (Akaike's Information Criteria) of model 2 is smaller than that of model 1. Accordingly, the model 2 is able to explain the 2-step decision making process of mode choice behavior more accurately than the model 1. From the view point of parameters'

Table 4 – Estimated Parameters of PLCS models

step	variables			model 1	model 2
step1	min. difference of GC	Acap	min. difference of AGC		-2.1978 (-7.226)
		Bcap	min. difference of BGC		-0.4529 (-2.205)
$Q_n(G_i)$	$A.I.$	Rcap	$A.I._1$	2.9373 (16.30)	2.8640 (15.94)
		Acap	$A.I._2$	3.5877 (18.92)	3.0242 (16.23)
		Bcap	$A.I._3$	2.0356 (13.22)	2.1545 (13.12)
		sel	$dif.A.I.$	3.3287 (8.694)	2.7151 (7.009)
step2	LOS	C	travel time (hour)	-1.2696 (-7.926)	-1.1041 (-6.489)
$P_n(i\mid G_0)$		C	travel cost (10,000yen/person)	-2.8694 (-5.161)	-2.1691 (-3.617)
	constant term	R	express train constant term	-0.7937 (-1.863)	-0.6604 (-1.419)
		A	automobile constant term	-2.2061 (-5.366)	-2.4536 (-5.209)
$\bar{\rho}^2$				0.550	0.568
AIC (Akaike's Information Criteria)				1597.56	1535.77
hit ratio of mode choice result (%)				81.4%	82.1%
hit ratio of alternative category (%)				55.2%	59.2%
value of time (yen/minute)				73.7	84.8
number of samples				1,611	1,611

* ( ) : statistical t-value
* Rcap : express train captive Acap : automobile captive Bcap : inter-city express bus captive
sel : selective
* C : common variable R: express train specific variable A: automobile specific variable

significance of explanatory variables, each variable except for express train constant term has significant t-value. Both difference of GC and $A.I.$ have impact on the "Alternative screening process (step1)" with statistical efficiency. LOS, such as travel time and cost, also has statistical efficiency in the utility functions. This result seems to prove that the assumption of decision making process of mode choice behavior of inter-regional travellers with non-business purpose in the study (figure 9) is reasonable and appropriate.

Sensitivity analysis is also done with the PLCS model (model2) to observe how $A.I.$ influent to the decision making process of mode choice behavior. For this analysis, these five assumptions are applied as follows:

A) Objective trip is from F city (representative point is H station) to N city (representative point is N station). Trip distance is about 160km.
B) Objective $A.I.$ is fluctuated from average 0 to +2.3. 0 is average amount of $A.I.$. +2.3 is relevant to maximum amount of $A.I.$. Because $A.I.$ is latent factor score which is according with standard normal distribution (average: 0, standard deviation: 1).
C) 0 is substituted to the non-objective $A.I.$ as the average amount of $A.I.$.
D) LOS data of above-mentioned travel is substituted to the model.
E) Constant terms are adjusted to the modal split share from F city to N city observed in INFS conducted in 2005.

The result of sensitivity analyses of each $A.I.$ are shown in Figure 11~13. When every $A.I.$ is 0, which is equivalent to the case of $dif.A.I.$ is equal to infinity, probability of selective category is 100%. On the other hand, probability of captive category of each mode and the estimated share are increasing in accordance with increasing $A.I.$ of each mode. This suggests that latent preference factors explained by $A.I.$ have strong impact to the decision making process of mode choice behavior of inter-regional travellers.

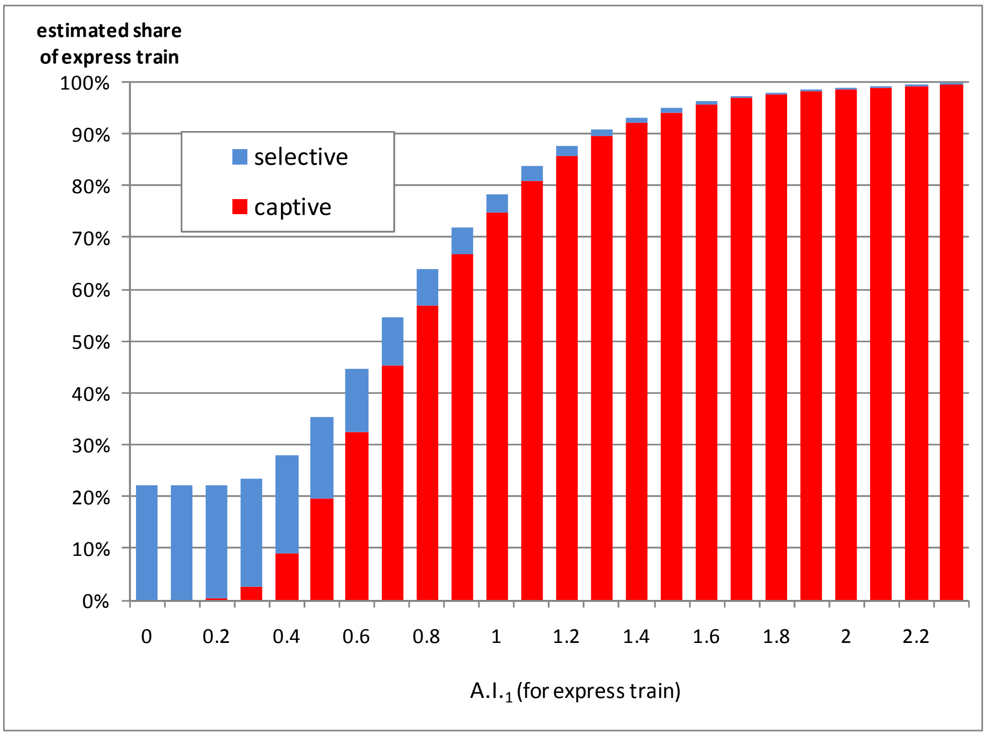

Figure 11 – Sensitivity of $A.I._{1}$ (for express train)

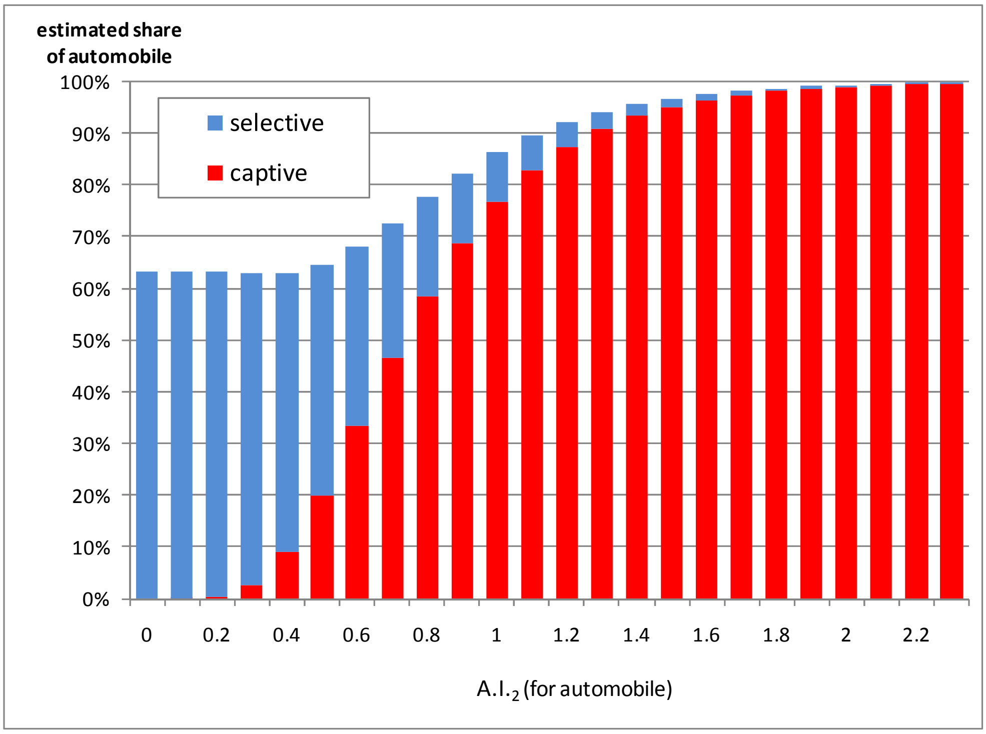

Figure 12 – Sensitivity of $A.I._{2}$ (for automobile)

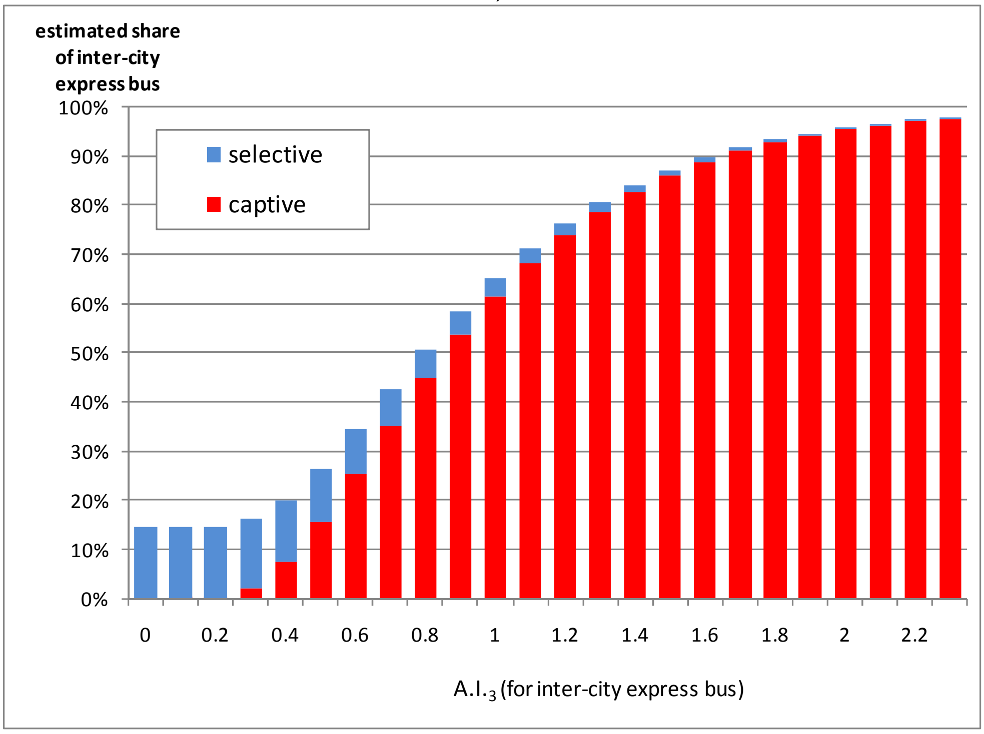

Figure 13 – Sensitivity of $A.I._3$ (for inter-city express bus)

4.3 Effect of Market Segmentation by the PLCS Model

Figure 14 and Figure 15 show the comparison of sensitivity between PLCS model (Table 4; model 2) and ordinary commonly-used disaggregate logit model (Table 2; model 2). As for both travel time and travel cost, PLCS model is more sensitive than ordinary logit model. This suggests that travellers belonging to the selective category choose transportation mode mainly by LOS. In other words, such travellers choose transportation mode more reasonably from the view point of analyst. And the market of inter-regional travellers with non-business purpose should be segmented into the mode captive categories and selective category in order to analyse mode choice behavior more accurately.

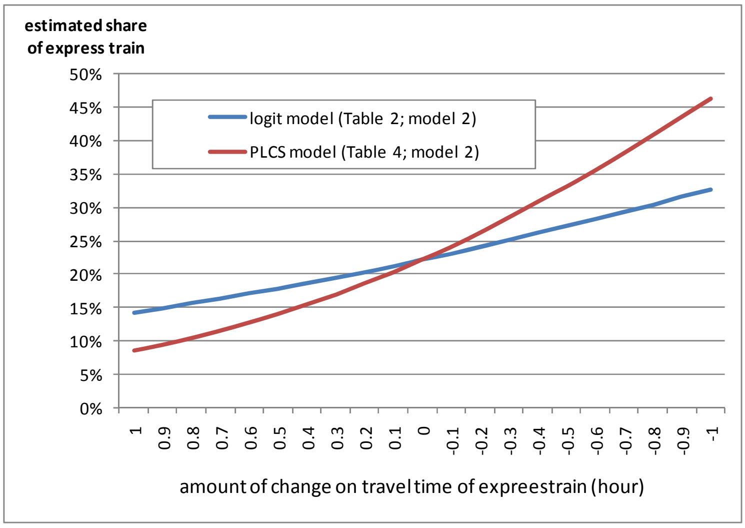

Figure 14 – Comparison of Sensitivity of PLCS Model and Logit Model (as for travel time)

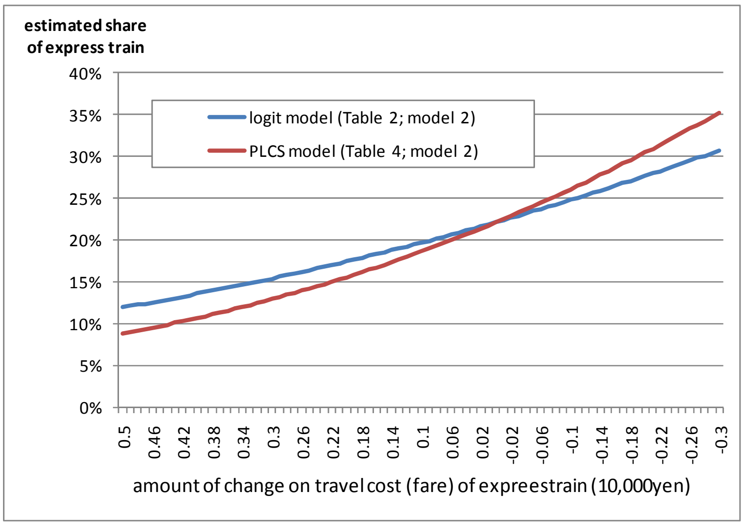

Figure 15 – Comparison of Sensitivity of PLCS Model and Logit Model (as for travel cost)

5 CONCLUSION REMARKS

The study focuses on mode choice behavior of inter-regional travellers with non-business purpose particularly from the view point of effectiveness of latent preference factors to the decision making process. Some aggregate and disaggregate analyses prove that latent preference factors have high impact on not only mode choice behavior, but also mode alternative screening process. This suggests that mode choice behavior should be treated as at least 2-step decision making process, "Mode alternative screening process" and "Determining process". However, commonly-used disaggregate logit model can describe only the "Determining process". Thus, the study proposes the PLCS (Parameterized Logit Captivity and Selectivity) model in order to describe the 2-step decision making process of mode choice behavior. Some sensitivity analyses suggest that the market of inter-regional travellers with non-business purpose should be segmented into the mode captive categories and selective category in case of analyses of mode choice behavior.

Finally, the study concludes

A) The share of mode captive usually keeps high sharing rate in every inter-regional mode user. Thus, mode choice behavior should be treated as at least 2-step decision making process, "Mode alternative screening process" and "Determining process".
B) The PLCS model proposed by the study, a kind of extended PLC model, is efficient to describe the 2-step decision making process of mode choice behavior.
C) In particular, latent preference factors, such as "potential preference for each transportation mode" and "magnitude of importance for intangible factors", greatly influent to the "Mode alternative screening process" suggested mainly by the PLCS model.
D) The market of inter-regional travellers with non-business purpose should be segmented into "mode captive categories" and "selective category" in order to analyse mode choice behavior more accurately.

REFERENCES

Okumura, M. and Tsukai, M. (2007) Air-Rail Inter-Model Network Design under Hub Capacity Constraint, Journal of Eastern Asia Society for Transportation Studies, Vol. 7, pp.180-194

Shibata, M., Muto, M. and Uchiyama, H. (2001) A Modal Split Model for Inter-Regional Travellers on Holidays with the Consideration of Intangible Factors, Journal of Eastern Asia Society for Transportation Studies, Vol.4, No.3, pp.301-313

Muto, M. and Uchiyama, H. (2000) A Study on Service Plan for Inter-city Express Train based on the Holiday Passenger Behavior, Infrastructure Planning Review, No.17, pp.745-750 (in Japanese)

Morikawa, T. and Sasaki, K. (1993) Discrete Choice Models with Latent Explanatory Variables using Subjective Data, Journal of Infrastructure Planning and Management, No.470 / 4-20, pp.115-124 (in Japanese)

Muto, M., Shibata, M., Hibino, N. and Uchiyama, H. (2004) A Study on Mode Choice Behavior for Inter-regional Travelers on Holidays Focused on Subjective Factors, Transport Policy Studies' Review, Vol.6, No.4, pp.2-11 (in Japanese)

Shibata, M. and Uchiyama, H. (2009) A Study on the Decision Process of Mode Choice Behavior in Inter-regional Trips, Infrastructure Planning Review, Vol.26, No.3, pp.457-468 (in Japanese)

Swait, J. and Ben-Akiva, M. (1987) Empirical Test of a Constrained Choice Discrete Model: Mode Choice in SAO PAULO, BRAZIL, Transportation Research-B, No.2, pp.103-115

Shibata, M., Muto, M., Tamura, K., and Li, G. (2006) An Analysis on Mode Choice Behavior of Inter-Local Regions Passengers on Public Trunk-line Service, Proceedings of Infrastructure Planning, Vol.33, CD-ROM (in Japanese)

Center for Spatial Information Science, the University of Tokyo Geocoding Tools & Utilities, http://newspat.csis.u-tokyo.ac.jp/geocode/ (in Japanese)

Ministry of Land, Infrastructure, Transport and Tourism (2009) Introduction of functions of NITAS, http://www.mlit.go.jp/seisakutokatsu/soukou/nitas/nitasout.pdf (in Japanese)

Ikeda, H. (1986) The Successive Category Method and the Application, The Japanese Journal of Ergonomics, Vol.22, No.4, pp.185-190 (in Japanese)

Manski, F. C. (1977) The Structure of Random Utility Models, Theory and Decision 8, pp.229-254, D. Reidel Publishing Company, Dordrecht-Holland