Characteristics of inter-regional passenger transportation demand by repeated web survey

Yushi SUGAHARA[1], Makoto TSUKAI[2] and Canh Xuan DO[3]

[1]Student Member, Graduate School of Engineering, Hiroshima University
(1-4-1, Kagamiyama, Higashi-hiroshima City, Japan)
E-mail: m176539@hiroshima-u.ac.jp
[2] Regular Member, Graduate School of Engineering, Hiroshima University
(1-4-1, Kagamiyama, Higashi-hiroshima City, Japan)
E-mail: mtukai@hiroshima-u.ac.jp
[3] Student Member, Graduate School of Engineering, Hiroshima University
(1-4-1, Kagamiyama, Higashi-hiroshima City, Japan)
E-mail: canh.doxuan@gmail.com

In Japan, the Ministry of Land, Infrastructure, Transport and Tourism has conducted the Net Passenger Transportation Survey (hereafter, called NPTS) since 1990, which gives valuable information about inter-regional passenger flows. However, NPTS has been criticized in several limitations. One of them is that this survey does not consider about seasonal differences or trip frequency of observed passengers. Moreover, seasonal differences and trip frequency are also important for supplying of accommodation capacity or transfer site design. Therefore, in order to overcome the above limitations, a periodical survey using web survey is essential.

By using repeated web survey data, this study employs a model to estimate inter-regional trip frequency and a model to predict inter-regional passenger flows. Trip frequency is separately estimated in each season by using negative binomial models. The modal choice model with two modes (e.g., rail and air) is estimated considering trip frequency weight.

Key Words: *trip frequency, modal choice, sample weight, inter-regional travel demand*

1. INTRODUCTION

For the network design, inter-regional passenger demand using domestic network is necessary for giving fundamental information to network planning. The Ministry of Land, Infrastructure, Transport and Tourism has been conducted Net Passenger Transportation Survey (hereafter, called NPTS) every five years since 1990 [1]. NPTS investigates inter-regional passenger flow for each transportation mode such as air, express train, ship, car and bus. Although NPTS is essential in the field of transportation planning, it has been criticized in the several limitations as follows.

NPTS is conducted in one weekday and one holiday in autumn. On the other hand, it is well known that inter-regional passenger flows are characterized by seasonally unstable, which is different from the intra-city trips in commuting (e.g., going to work or school). Unfortunately, NPTS cannot provide enough information to accommodation planning or local transport capacity planning since it does not include the seasonal variation of inter-regional passenger flows. Furthermore, publishment of NPTS usually takes one or two years in the post-survey processing, since the estimation of expansion coefficient for each data sample costs much time.

Another limitation of NPTS comes from survey method. NPTS is an on-trip survey which is mainly conducted by directly distributing questionnaire to travelers while the home-based survey is implemented through sending questionnaires by mail or by email. Therefore, in order to make a useable data for forecasting inter-regional passenger flow from on-trip survey, it is necessary to calculate an expansion coefficient for each sample, which is equivalent to the sampling rate of each mode or route. Such the procedure would be effective if the sampling rate of each mode is relatively high. However, the sampling rate for rail passengers is considerably low while that of air passengers is relatively high in NPTS, so that the reliability in rail trip is often questioned. Such the difference in sample reliability causes a problem in the modal choice model between air and rail.

Moreover, NPTS cannot give any information of non-trip makers. In case of forecasting traffic demand in the future, paying attention to such the latent passengers as a novel tourism target is necessary.

In order to tackle the problems mentioned above, a home-based survey with a long-term continuous observation should be considered. However, performing a long-term continuous observation would be quite difficult since it takes long time and high expense in all survey stages such as distributing and collecting questionnaire sheets and post-survey data processing. Instead, conducting a web survey would be more efficient.

Web survey has two advantages. First, web survey can obtain a large number of samples in a short time. The second one is that the sample of non-trip makers are also available since web survey is one type of home-based surveys. However, web survey still has several disadvantages, one of them is response bias which is often mentioned. Fortunately, correction methods by using the propensity scores method have already been proposed in past studies. On the other hand, in repeated web surveys, exploring characteristics of inter-regional passenger flows or considering sample weighting method has not been studied so far. Therefore, in order to complement NPTS, home-based web survey would be an alternative solution.

This study uses a series of web survey conducted every three months and repeated four times to cover a whole year from April 2015 to March 2016. This survey includes questions about the three-month-period travel history which is made before each wave of our survey. By assigning unique identity code to each respondent in order to distinguish those who participated in all four waves of our web surveys, panel inter-regional travel demand data is collected. By utilizing this data, seasonal variations in travel demand characteristics is clarified. This study employ two model: (1) one model is estimate the trip frequency of inter-regional passenger flow; and (2) another is modal choice model with the sample weight by trip frequency.

This paper is organized into following sections. The next section shows the literature review. Section 3 illustrates two models employed in this study. Section 4 summarizes the web survey. Section 5 presents the result of estimating trip frequency. Section 6 discusses the result of modal choice model. Section 7 is the conclusion.

2. LITERATURE REVIEW

(1) Net Passenger Transportation Survey (NPTS)

NPTS is designed to get domestic inter-regional passenger flows at certain sections. This survey also

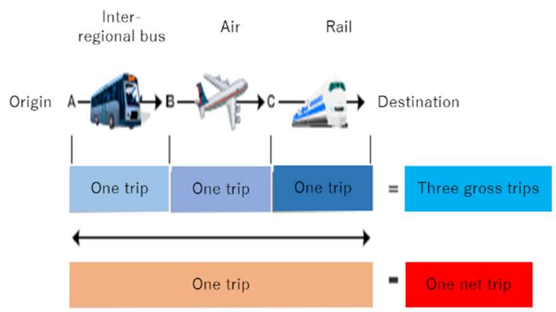

Figure 2-1 An outline of net trip and gross trip

collects individual characteristics of travelers who use airlines, express train, ship, inter-regional bus and car. NPTS records a whole trip from the first departure area (origin) to the last arrival area (destination), including some transfers at train stations, airports, ship ports, bus stops or junctions. A trip discussed above is called "net trip", in contrast, a trip separately recorded for each transportation mode is called "gross trip". **Figure 2-1** describes the concepts of a net trip and a gross trip.

(2) Studies in inter-regional net passenger traffic

Isono [2] pointed out that in the field of transportation planning in tourism policy, passenger flow data with the seasonal variation in demand is required to know the true origins and destinations of passengers since inter-regional passenger demand greatly fluctuated over seasons. In his study, the number of monthly OD pair trips among prefecture with trip purpose and a representative mode were estimated by integrating several existing data. Furthermore, in-depth analysis on tourism policy supported or tested by passengers' travel demand characteristics is very crucial.

Okumura [3] summarized issues of NPTS. In case of the intra-city daily traffic survey, focusing on travel demand in an average day is adequate because trip pattern and frequency do not change in daily. However, in inter-regional passenger flows survey, collecting for an average demand is not appropriate, since inter-regional travel flows are greatly affected by consecutive holidays, seasons or even weather. Therefore, Okumura proposed using the web survey method and other approaches to collect trip samples in various seasons, which would be better than collecting trip samples by an on-trip questionnaire survey.

Furthermore, Okumura [3] also referred to the set problem in expansion coefficients in NPTS. Basically, the expansion coefficients in NPTS depend on the ratio between the number of samples and the volume of demand at the surveyed link. If an expansion coefficient is given by considering simply the sample ratio, it is impossible to correct the deviation of trip

Table 3-1 Types of negative binomial model

Type	Assumption of variance
NEGBIN I	$V[Y] = E[Y] \times (1 + \alpha)$
NEGBIN II	$V[Y] = E[Y] \times (1 + \alpha \times E[Y])$

purposes or travelers' attributes. In order to correct them, another approach to computing the expansion coefficient is necessary.

3. MODEL SPECIFICATIONS

(1) Trip frequency model

Here in this study, a count model with negative binomial distribution is employed and specified as in equation (1).

$$f_{(X=y)} = \frac{e^{-\lambda}\lambda^{y}}{y!} \quad (1)$$

where λ is a parameter following to a gamma distribution in equation (2) and y is a count number as well as a dependent variable representing the trip frequency.

$$g(\lambda) = \frac{c^{r}}{\Gamma(r)} e^{-c\lambda}\lambda^{r-1} \quad (2)$$

where r is an overdispersion parameter represented to the variance of distribution, and $\Gamma(.)$ illustrates the gamma function.

There are shown two types of negative binomial models regarding variance definition as shown in **Table 3-1**with α is a constant parameter. In this study, the variance of trip frequency is specified in NEGBIN II. In NEGBIN II, probability density function is specified in equation (3)

$$f_{(X=y)} = \frac{\Gamma(y+r)}{y!\,\Gamma(r)} \left(\frac{r}{\mu+r}\right)^{r} \left(\frac{\mu}{\mu+r}\right)^{y} \quad (3)$$

where μ is a mean of y with the property $\mu=E[Y]$

(2) Modal choice model

In a trip, the traveler uses one of several transportation modes such as a private car, railway, airline and so on. In transport demand analysis, modal choice mechanism could be explained by utility maximization. The explanatory variables in a utility function could be the level of transportation service, traveler's attributes and so on. For modal choice analysis, the probability of choosing mode m is specified in equation (4) as follows.

$$P_m = \frac{exp(V_m)}{\sum_j exp(V_j)} \quad (4)$$

where m is a representative mode in a net trip, P_m is a probability of selecting mode m and V_m is a utility function of mode m.

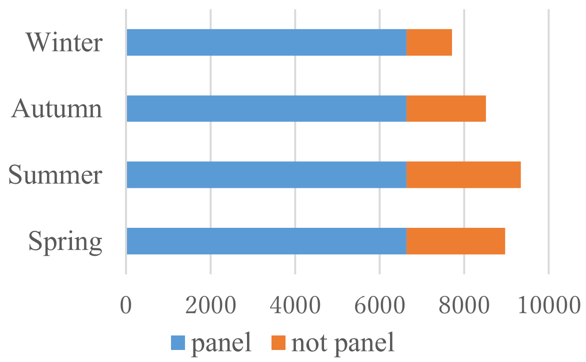

Figure 4-1 The number of respondents

4. DATA AGGREGATION

A questionnaire of web survey in this study was designed to observe the seasonal difference in inter-regional passenger trip demand. The survey was conducted four times in August and October in 2015 and January and April in 2016 by a web survey monitor named Intage Co. Ltd, a company for marketing research.

The respondent, who voluntarily applied to our survey, answered a questionnaire at home (i.e., home-based survey) using internet service. Each was given a unique identification number (ID number) at the first survey in August 2015. Therefore, in the following surveys, respondents could be recognized by their ID. Based on these ID, respondents are divided into two groups including panel group or non-panel group. Panel group includes respondents who participated in all four waves of our surveys, and vice versa. **Figure 4-1** shows the share of panel group and non-panel group. The number of respondents in the panel group is 6,634 and those of non-panel is 5,676.

In our survey, each record includes the respondent attributes and inter-regional trips made in a designed period (i.e., latest three months before each wave of our survey). The travel purposes are categorized in business, sightseeing, private and others. Commuting trips are excluded from the survey since these trips are usually made as one of daily activities and these trips' frequency is often more than once in a day. Respondent's attributes asked in this survey are sexuality, age, occupation, income, family, home address and so on. Each respondent could answer up to three trips with different destinations including trip frequency made in three months. In this study, we used those samples which made by the panel group. **Table 4-1** shows a summary of web survey. **Figure 4-2** shows share of trip frequency aggregated in the following six categories such as no trip, 1 time, 2 times, 3 times, 4 to 9 times and over 9 in each season. The first wave of our survey in spring, does not consist of samples from with no trip makers because in

Table 4-1 Summary of web survey

Survey content	Description
Survey period and survey time	Spring : August 2015 Summer : October 2015 Autumn : January 2016 Winter : April 2016
Respondents	Aged at 20 and over who live in Tokyo, Chiba, Saitama and Kanagawa
Number of respondents	34,534 in all seasons
Trip purposes	Business, sightseeing, private and others (excluding commuting)
Items (individual attributes)	Gender, age, occupation, income, family, hometown and so on
Items (trip attributes)	Frequency, departure area, visited area, all the modes used in trip and so on

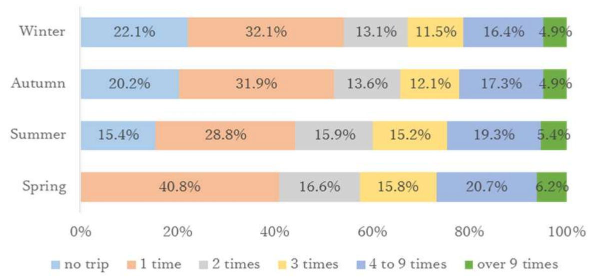

Figure 4-2 Trip frequency by seasons

that wave only samples of trip makers were collected by the web survey monitor. As shown in to **Figure 4-2**, travelers tend to make more trips in summer than other seasons.

Regarding data comparison, sample observation characteristics of each survey should be considered. One of NPTS limitations is that a question about the trip frequency of the observed trip is not included in the questionnaire. Since the respondents of NPTS is the on-trip passengers, the probability to observe frequent trip makers would be higher, while that of infrequent trip makers would be lower. Therefore, the probability of each observed trip is proportional to trip frequency of that trip. In order to consider trip frequency, a sample weight is added to each trip based on the probability of observed sample as defined in equation (5).

$$w_s = \frac{t_s/90_{(days)}}{\sum_s (t_s/90_{(days)})} \times N = \frac{t_s}{\sum_s t_s} \times N \quad (5)$$

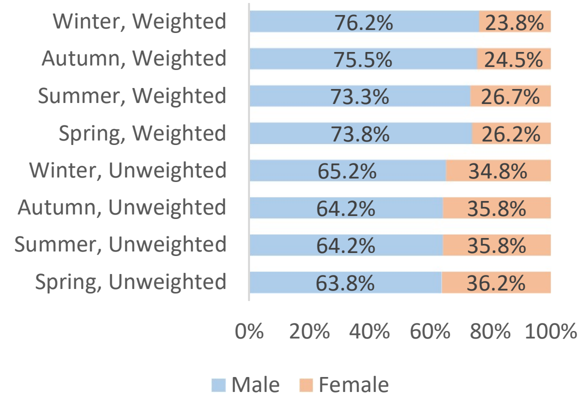

Figure 4-3 Gender share by seasons with and with out sample weight

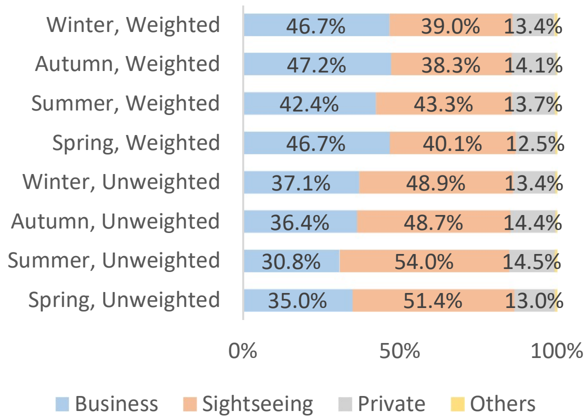

Figure 4-4 Trip share by trip purpose and seasons with and with out sample weight

where N is a number of samples, w_s is a sample weight for a sample s and t_s is the trip frequency of each recorded trip in sample s.

Figure 4-3 and **Figure 4-4** respectively illustrates shares of genders and trip purposes over four seasons with the trip weight (weighted samples) and without sample weight (unweighted samples). If a weighted share of an attributes is larger than that of unweighted, the respondents who have the attributes make more trips than the others, in other words, per capita trip frequency of the respondents is higher than that with other attributes. For example, in **Figure 4-3**, the share of males is larger in weighted samples in all seasons. Therefore, males tend to make more trips per capita than females. In **Figure 4-4**, share of business trips is larger in weighted samples in all seasons, which means that business trips are made per capita more frequency than the other purposes.

5. ANALYSIS ON TRIP FREQUENCY

(1) Trip frequency distribution

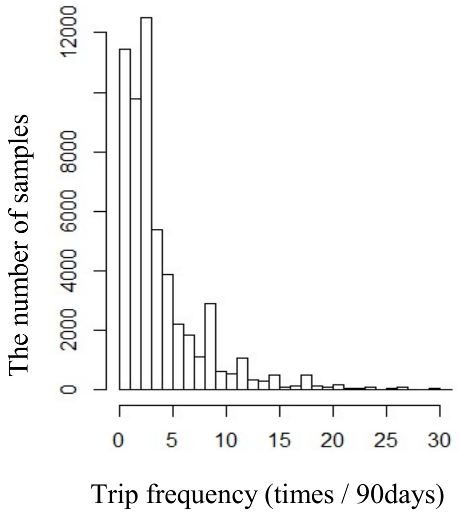

Figure 5-1 Sample distribution of trip frequency in all seasons

Figure 5-1 demonstrates a histogram of trip frequency in the latest three months. In this study, the negative binomial model type II specified in equation (3) is used to estimate the trip frequency.

$$\boldsymbol{\mu} = \exp(\boldsymbol{x\beta}) = \boldsymbol{E}[y] \tag{6}$$

where μ is an expected value, y is the trip frequency in each season (objective variable), x is dummy variables about a respondent attribute (explanatory variable vectors) and β is an estimated parameter vector of x in each season

The parameter vector β is estimated by log-likelihood maximization. The trip frequency is increased if estimated parameter is positive and vice versa.

For convenience in interpretation, parameters in the model are categorized. By the range of estimated parameters between the largest and the smallest in each category, the model gives the impact on trip frequency.

R^2_{DEV} is the model fit index specified in equation (7).

$$R^2_{DEV} = 1 - \frac{\sum_i \left\{ (\hat{\mu}_i + r) * log\left(\frac{\hat{\mu}_i + r}{y_i + r}\right) + y_{wi}\, log\left(\frac{y_i}{\hat{\mu}_i}\right) \right\}}{\sum_i \left\{ (\hat{\mu}_i + r) * log\left(\frac{\bar{y} + r}{\hat{\mu}_i + r}\right) + y_{wi}\, log\left(\frac{y_i}{\bar{y}}\right) \right\}} \tag{7}$$

where y_i is observed trip frequency on individual i, $\bar{y}$ is average of y_i and $\hat{\mu}_i$ is estimated trip frequency on individual i.

(2) Trip frequency model in all seasons

In this section, trip samples from the panel group are used to estimate trip frequencies.

Table 5-1 Estimated parameters in trip frequency model

Category	item	Estimate	
	Constant	0.080	**
Season	Spring	0.137	***
	Summer	0.047	***
	Autumn		
	Winter	-0.027	*
Gender	Male	0.222	***
	Female		
Age segments	20-39	-0.117	***
	40-49		
	50-59	-0.027	*
	Over 60	-0.090	***
Occupation	Manager	0.249	***
	Salaryman	0.133	***
	Woreker		
	Student	0.282	***
	Part time	-0.098	***
	Non-woreker	-0.079	***
	Other	0.099	*
Income	Less than million	0.204	***
	1-5 million	-0.036	**
	5-10 million		
	More than 10 million	0.111	***
	No answer	0.052	***
Marry	Married	0.099	***
	Non-maried		
Child	No child	0.038	
	One		
	Two	0.041	**
	Over three	0.207	***
Family	Single	0.112	***
	With partner	-0.087	***
	Partner and child		
	With parents	-0.027	
	Other	0.019	
Hometown	Gunma	0.336	***
	Saitama	0.033	*
	Chiba		
	Tokyo	0.033	
	Kanagawa	0.080	***
	r	3.198	***
	-2 log L	1.553.E+05	
	R^2_{DEV}	0.075	
	Number of samples	46832	

Significant level : '***' 0.1% '**' 1% '*' 5% '.' 10%

The result of model estimation is shown in **Table 5-1**. R^2_{DEV} shows model fit to data, considering the comparison between estimated value and observed dependents variable. Therefore, the estimated model does not fit. In **Table 5-1**, the estimated parameter of the variable of spring the highest in the category of seasons, followed by the estimated parameter of summer. This is because non-trip samples are not observed in the first wave of our survey (i.e., spring season). This result indicates that trip frequency is significantly higher in summer due to long vacation.

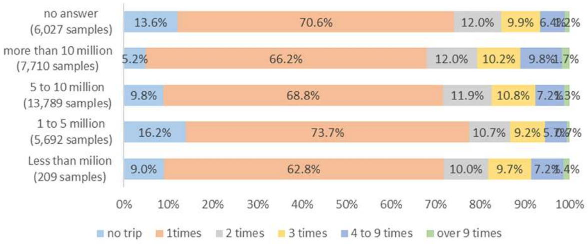

Figure 5-2 Share of trip frequency by each income segments

Table 5-2 The range of maximum and minimum

Season / Category	all Range	all Range rank
Season	0.164	7
Gender	0.222	4
Age segments	0.117	8
Occupation	0.380	1
Income	0.240	3
Marry	0.099	9
Child	0.207	5
Family	0.199	6
Hometown	0.336	2

Generally, most of variables of travelers' attribute have influenced on the trip frequency. To be more detail, males, managers, salarymen and high income people tend to be frequent passengers. This result might be consistent with the higher share in business trips in Japan. Students and those who have income less than millions tend to make more trip. However, the implication of the above fact is a bit difficult to understand. One possibility behind this fact is that such people have enough time to make trips but they would not spend much money for their trips, thus low cost carrier (LCC) of travel modes would be preferred. **Figure 5-2** shows the aggregation of trip frequency in latest three months on each segment of income for a whole year (i.e., all seasons). Traveler having income less than million tend to make more trip than those having income from 1 to 10 million and un-answered income (i.e., not answered about their income).

Table 5-2 shows the range between the lowest to the highest value in each category, implying the degree of influence on the trip frequency. The highest value is shown in occupation, followed by hometown, income and gender. On the other hand, age or marry are relatively small.

Since samples used in this study are extracted from panel group, the share of individual attributes is equal in all seasons. The model estimated for each seasons is shown in following section.

Table 5-3 Estimated parameters in trip frequency model for each season

Category	Season / item	Spring Estimate		Summer Estimate		Autumn Estimate		Winter Estimate	
	Constant	0.297	***	0.197	***	-0.088		0.035	
Gender	Male	0.196	***	0.229	***	0.226	***	0.240	***
	Female								
Age segments	20-39	-0.075	***	-0.137	***	-0.124	***	-0.132	***
	40-49								
	50-59	-0.004		-0.034		-0.067	**	0.001	
	Over 60	-0.031		-0.186	***	-0.092	**	-0.050	
Occupation	Manager	0.183	***	0.253	***	0.357	***	0.214	***
	Salaryman	0.090	***	0.118	***	0.193	***	0.143	***
	Woreker								
	Student	0.661	***	0.156		0.092		-0.069	
	Part time	-0.113	**	-0.118	*	-0.086	.	-0.074	
	Non-woreker	-0.089	**	-0.076	*	-0.028		-0.127	**
	Other	0.049		0.010		0.471	***	-0.221	.
Income	Less than million	0.070		0.187		0.286	*	0.323	*
	1-5 million	-0.050	*	-0.008		-0.021		-0.069	*
	5-10 million								
	More than 10 million	0.107	***	0.029		0.182	***	0.130	***
	No answer	0.027		0.016		0.032		0.140	***
Marry	Married	0.073	*	0.105	**	0.165	***	0.065	
	Non-maried								
Child	No child	0.053		0.013		0.004		0.085	
	One								
	Two	0.007		0.032		0.047		0.081	*
	Over three	0.275	***	0.097	*	0.297	***	0.130	*
Family	Single	0.088	.	0.082		0.219	***	0.070	
	With partner	-0.087	.	-0.055		-0.059		-0.158	**
	Partner and child								
	With parents	-0.041		-0.037		0.126	*	-0.150	*
	Other	-0.017		0.084		0.048		-0.038	
Hometown	Gunma	0.426	***	0.416	***	0.212	.	0.196	
	Saitama	0.010		0.047		0.075	*	-0.003	
	Chiba								
	Tokyo	-0.009		0.011		0.063	*	0.079	**
	Kanagawa	0.067	**	0.057	*	0.081	*	0.118	***
	r	4.354	***	3.153	***	2.841	***	2.612	***
	-2 log L	4.196.E+04		4.027.E+04		3.703.E+04		3.585.E+04	
	R^2_{DEV}	0.077		0.067		0.089		0.070	
	Number of samples	12499		12091		11280		10962	

Significant level : '***' 0.1% '**' 1% '*' 5% '.' 10%

(3) Trip frepuency model for each season

Table 5-3 shows the result of trip frequency model for each season. Through the year, items such as male, manager and salaryman are estimated with positive and significant. This result shows that business trips are frequent than other purposes, shown in **Figure 4-4**. Since, 20-39 is estimated with negative and significant, their frequency is significantly low.

Table 5-4 shows the ranges estimated by each season. Occupation has the largest range in all seasons, while marry has the smallest range in almost all seasons. Hometown has larger range in spring and summer, smaller range in autumn and winter. Income has smaller range in spring and summer, but larger range in autumn and winter. Child has relatively larger range in spring and autumn. Comparing in seasonal tendency, the rank in category is quite similar in spring and autumn, except income and hometown which are swapped. As same as spring / autumn, summer and winter in also similar range rank except income and hometown, which is swapped in the couple of categories.

Table 5-4 The range of maximum and minimum

Season	Spring		Summer		Autumn		Winter	
Category	Range	Rank	Range	Rank	Range	Rank	Range	Rank
Gender	0.196	4	0.229	3	0.226	5	0.240	3
Age segments	0.075	7	0.186	5	0.124	8	0.133	6
Profession	0.774	1	0.371	2	0.557	1	0.435	1
Income	0.157	6	0.195	4	0.307	2	0.392	2
Marry	0.073	8	0.105	7	0.165	7	0.065	8
Child	0.275	3	0.097	8	0.297	3	0.130	7
Family	0.174	5	0.138	6	0.278	4	0.228	4
Hometown	0.435	2	0.416	1	0.212	6	0.199	5

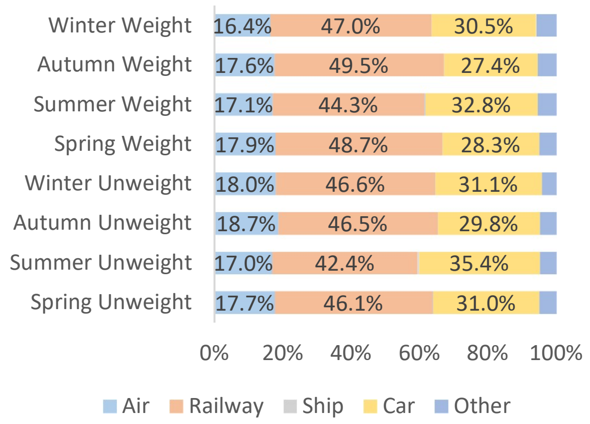

Figure 6-1 Main transportation mode

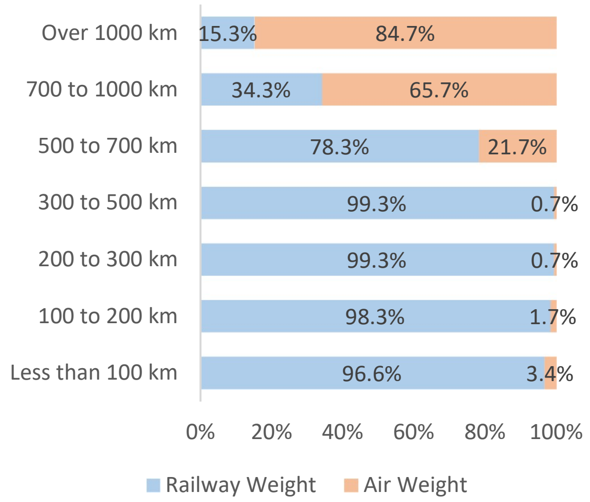

Figure 6-2 Length of trip

6. MODAL CHOICE MODEL

(1) Subject of transportation mode

The web survey contains several questionnaire items about transportation modes (air, railway, ship, car and others) on net passenger trip. **Figure 6-1** shows the share of the representative mode. As mentioned in Section 4, for comparison, data aggregation is made for both unweighted and weighted samples. In all seasons the share of railway and air are slightly larger in case of weighted samples than in case of unweighted samples, while the share of car is bit smaller in case of weighted samples than in case of unweighted samples. If the weighted share of an attribute is larger than unweighted share, per capita trip frequency with the attribute is relatively large and vice versa. Therefore, rail and air are selected as travel modes since they would attract many passengers in long distance trips. Figure 6-2 shows that, in over 500 km trip range, the longer trip is, the higher share of air is. In the following section, modal choice model is estimated in which level of service (LOS) of both modes is calculated by using NITAS [5].

(2) Modal choice model

The probability of choosing mode m is formulated in equation (4). In this study, two selected alternatives of modes are air and railway. Therefore, the utility function for air or rail is formulated in equation (8) and (9).

$$V_i^R = \sum_k \beta_k X_{ki}^R + \sum_l \beta_l X_{li} \tag{8}$$

$$V_i^A = \sum_k \beta_k X_{ki}^A \tag{9}$$

where $X_{ki}^{R/A}$ are common variables in LOS for both alternatives, travel time and travel cost, X_{li} are alternative specific variables in railway, β_k and β_l are parameters to be estimated. The parameters are estimated by maximum likelihood method.

Log of likelihood function with sample weight is specified in equation (8).

$$log(L) = \sum_s w_s \times \log(P_s(i)) \tag{10}$$

where s is a sample, w_s is a sample weight for s and $P_s(i)$ is the probability of choosing mode i by each sample s. Sample weight w_s is given by referring to trip frequency of each destination record.

(3) The result of model estimation

The following two models with and without the sample weight in equation (10) are estimated. The difference in such treatment may cause the different tendency in the significance level of each explanatory variable.

Table 6-1 shows the comparison of model estimation between NPTS, weighted and unweighted in autumn since NPTS is conducted in autumn. NPTS model is estimated by equation (10). However, the daily expansion coefficient is used as a sample weight instead the weight w_s in web survey. Significance and likelihood ratio are larger in weighted than that of unweighted. Therefore, weighted model is better than unweighted model to estimate modal choice. Comparing weighted with NPTS, likelihood ratio of web survey is lower. Then, we reestimate the web survey model, focusing on some trip samples which trip distance is over 300 km.

Table 6-1 The result of modal choice model in autumn

		NPTS		Weighted		Unweighted	
	item	estimate		estimate		estimate	
	Constant (railway use)	-1.886	**	-0.301	*	-0.949	**
	time (/60 min)	-0.200	**	-0.112	**	-0.156	**
	fare (/1,000 yen)	-0.058	**	-0.032	**	-0.029	**
Individual attribute	Male	-0.499	**	-0.659	**	-0.580	**
	20 to 39	0.863	**	0.451	**	0.262	*
	50 to 59	0.147	**	0.348	**	0.229	*
	Over 60	0.000		-0.021		0.102	
	Manager	0.875	**	-0.273	*	-0.118	
	Salaryman	-0.273	**	-0.443	**	-0.034	
Trip attribute	Business	0.527	**	0.542	**	0.424	**
	Private	0.539	**	0.522	**	0.142	
	Other	0.824	**	-0.096		0.645	
	Likelihood ratio	0.585		0.343		0.323	
	Fixed likelihood ratio	0.585		0.340		0.320	
	Amount of train trip	156047		4197		3595	
	Amount of air trip	31699		1498		1447	
	Time value (yen/h)	3431	**	3499	**	5440	**

Significant level : ‘***’ 0.1% ‘**’ 1% ‘*’ 5% ‘.’ 10%

Table 6-2 shows the result of modal choice model estimation, focusing on some trip samples which trip distance is over 300 km. Comparing the model focusing on trip distance and the model not focusing on trip distance, in both NPTS and weighted, likelihood ratio is lower than the model focusing on trip distance. Therefore, we estimate modal choice model by the weighted method, not focusing trip distance.

Table 6-3 shows the result of modal choice model in each season, by using panel data. Comparing the pooled data, likelihood ratio is lower than NPTS model. The constant parameter for railway is negative and significant in all seasons. Therefore, the unspecified utility for railway is lower than that for air. Parameters in travel time and travel fare/cost are negative and significant in all seasons. This is consistent with the normal expectation or the conventional condition in modal choice model. In case of gender, parameters in all seasons are negative and significant, therefore male prefers air. Those who are 20 to 29 tend to choose railway except winter. Railway is preferred by those who are 40 to 49 rather than those of the age group from 50 to 59, in summer, and vice versa. In terms of occupation, manager and salaryman tend to choose air except in winter. Therefore, air is likely preferred by people who get more income. In terms of purpose, railway is preferred except in winter.

Focusing on the value range in each explanatory variable category, constant has less influent effect on modal choice in all seasons. Gender has the highest impact on modal choice through the year.

Table 6-2 Modal choice model for over 300 km trip

		NPTS		Weighted	
	item	estimate		estimate	
	Constant (railway use)	-0.157	**	-0.294	*
	time (/60 min)	-0.131	**	-0.077	**
	fare (/1,000 yen)	-0.055	**	-0.038	**
Individual attribute	Male	-0.596	**	-0.098	
	20 to 39	0.586	**	-0.235	*
	50 to 59	0.075	**	-0.080	
	Over 60	0.066	*	-0.486	**
	Manager	0.570	**	-0.473	**
	Salaryman	-0.750	**	-0.270	*
Trip attribute	Business	-0.151	**	0.214	*
	Private	-0.302	**	-0.158	
	Other	0.287	**	3.697	.
	Likelihood ratio	0.384		0.194	
	Fixed likelihood ratio	0.384		0.190	
	Amount of train trip	90736		2658	
	Amount of air trip	31597		1459	
	Time value (yen/h)	2397		2018	

Significant level : ‘***’ 0.1% ‘**’ 1% ‘*’ 5% ‘.’ 10%

Table 6-3 The result of modal choice model in each season

		Web spring		Web summer		Web autumn		Web winter		NPTS	
	item	estimate		estimate		estimate		estimate		estimate	
	Constant (railway use)	-0.351	**	-0.483	**	-0.301	*	-0.561	**	-1.886	**
	Time (/60 min)	-0.149	**	-0.153	**	-0.112	**	-0.151	**	-0.200	**
	Fare (/1,000yen)	-0.026	**	-0.029	**	-0.032	**	-0.032	**	-0.058	**
Individual attribute	Male	-0.999	**	-0.840	**	-0.659	**	-0.511	**	-0.499	**
	20 to 39	0.284	**	0.026		0.451	**	0.291	**	0.863	**
	50 to 59	-0.005		-0.261	**	0.348	**	-0.087		0.147	**
	Over 60	0.020		-0.047		-0.021		0.451	**	0.000	
	Manager	-0.183	.	-0.294	**	-0.273	*	0.001		0.875	**
	Salaryman	-0.377	**	-0.238	*	-0.443	**	0.108		-0.273	**
Trip attribute	Business	0.714	**	0.706	**	0.542	**	-0.072		0.527	**
	Private	0.224	.	0.317	**	0.522	**	-0.254	*	0.539	**
	Other	0.548		0.567		-0.096		0.046		0.824	**
	Likelihood ratio	0.346		0.339		0.343		0.370		0.585	
	Fixed likelihood ratio	0.343		0.336		0.340		0.367		0.585	
	Amount of train trip	5025		4141		4197		3759		156047	
	Amount of air trip	1846		1594		1489		1312		31699	
	Time value (yen/h)	5783		5275		3499		4770		3431	

Significant level : ‘***’ 0.1% ‘**’ 1% ‘*’ 5% ‘.’ 10%

Occupation and purpose have less influent impact on modal choice in winter than other seasons. Among four seasons, winter shows bit different characteristics from others.

As shown in **Table 6-1 to 6-3**, an amount of train and air trips are calculated by multiplication of the number of trips and, the expand coefficient in NPTS or the sample weight in our web survey.

7. CONCLUSION

(1) Conclusion

This study pointed out the limitations in the current inter-regional passenger survey (NPTS), then tried to use web survey. Unlike NPTS, which is conducted in every five years, our web survey can collect the sample and surveyed data very quickly. Also, in order to clarify the difference in the inter-regional passenger demand, the web survey have conducted four times to look back the trips record over the latest three months. Then the trip frequency was analyzed by using the negative binomial model. Finally, a sample weight by trip frequency was given to the modal choice model, which is compared with the unweighted modal choice model.

From model results in trip frequency in Section 5, the seasonal demand change was significant, and the difference in the contribution of individual attributes was clarified. Since the dataset for modeling is limited to panel data, the seasonal difference is purely caused by the seasonal difference in passengers' difference in the trip generation by identical attributes.

In the weighted modal choice model in Section 6, the likelihood ratio is higher than that of the unweighted model, which means the fitness of the weighted is relative good. However, comparing with the NPTS model, this value is smaller, therefore, a further improvement of the weighted model is required.

(2) Future work

In the samples used in this study, the percentage of respondents who did not travel during the past three months is about 20% in summer, autumn, and winter. As a characteristic of our web survey, no-trip respondents may hesitate to join the survey. In other words, such the people directly results in non-observation. For the same reason, those who did not travel in three months may not continue to join the survey. Therefore, the number of samples in the no-trip group may be estimated in upward. Increasing the number of no-trip observations by changing the sample collection approach of the web survey may improve the fitness of the trip frequency model. Another possibility to improve the model fit is using the propensity score method. Kitahara [4] tried to model the trip frequency and trip interval using a web survey data. However, web survey is not efficient or has an another problem in sample collection method because home based web survey attracts the respondents who are more interesting in such this type of survey, so that not only the individual attribute but also the trip frequency would be biased. Kitahara used the propensity score method to reduce the bias in the web sample deviated from NPTS sample. Although the likelihood ratio of the propensity score model was relatively high, it was not possible to give sufficient correction effect on the trip frequency model. Such the disappointing result would be caused by trip frequency in a respondents' subgroup made by referring to the propensity score. As a future task, it is necessary to match the individual attributes share of samples with that of reference statistics.

In the weighted modal choice model, the likelihood ratio is lower than the NPTS model. It is important to estimate the simultaneous structure in modal choice behavior, destination choice behavior and the trip frequency choice. As the beginning, modal choice model alternative can be expanded to include bus, car or both.

REFERENCES

1) 国土交通省：全国幹線旅客純流動調査（NPTS），1990~2015
2) 磯野 文暁，ほか：季節変動を考慮した月別全国幹線旅客純流動データの推計，第39回土木計画学研究発表論文(347)，pp.1-4，2009
3) 奥村 誠：全国幹線旅客純流動調査の意義と課題，運輸と経済，第72 巻，pp.21-30，2012
4) 北原 聡子，ほか：Web 調査を用いたトリップ発生回数に関する研究，広島大学工学部卒業論文，平成24年度
5) National Integrated Transport Analysis System (NITAS) ： 国土交通省

(Received April 28, 2017)