Improvement of Interregional Passenger Demand Model by Propensity Score Method

Yoshiharu KATAGAWA[a], Makoto TSUKAI[b]

[a,b] *Graduate School of Engneering, Hiroshima University, Hiroshima, 739-8527, Japan*
[a]*E-mail: m126011@hiroshima-u.ac.jp*
[b]*E-mail: mtukai@hiroshima-u.ac.jp*

Abstract: National design about the rapid interregional passenger transportation system is critical for stimulating economic activities in Asian countries. However, the interregional demand forecasting including modal split between High-Speed-Rail and Airline based on the observed modal choice behavior faces the problems in imbalanced passenger observation among the modes. Such the imbalance would decrease the model performance in terms of reproductively. This study proposes a novel procedure to improve the demand forecasting about the regional level, based on the statistical modeling. For this purpose, propensity score can be utilized to give an adequate weight of each sample. The proposed sample weighting method based on the first principal component can improve the reproductivity of modal split.

Keywords: Principal Component Analysis, Aggregated Forecast, HSR, Airway

1. INTRODUCTION

1.1 Interregional Passenger Demand Forecasting

In Asian countries, national transportation design including the rapid interregional passenger transportation system is critical for stimulating economic activities. Therefore, the interregional demand forecasting including modal split between High-Speed-Rail (HSR) and Airline based on the observed modal choice behavior should be integrated to achieve the seamless transfer among them.

Up to now, many studies in interregional modal-choice are accumulated. Fu *et al.* (2012) estimated an inter-regional modal-choice model for air and rail passenger by using aggregated OD market data. And using the estimated model, they estimated the effects of introducing HSR. As a result, introduction of HSR from Tokyo-Osaka would almost perfectly exclude the airway demand.

Gong *et al.* (2011) conducted survey to clarify the leisure trip characteristics to the sightseeing spots characterized by the landmark, natural beauty and cultural heritage sites. The factors which have an effect on mode and destination choice were analyzed including passenger’s socio-demographic characteristics, trip purpose, conditions of local access, subjective modal preferences with tourist management policies. From the binary logit model analysis, the key factors to affect travel mode choice were travel scale, frequency, per capita daily consumption, age, private cars possession and attractiveness.

Kato *et al.* (2011) estimated the following three models: a modal choice, trip distribution, and tip generation/attraction using the interregional net passenger traffic survey in Japan. These models can give the moderate fitting to the data, but several explanatory variables cannot be set in the model due to the violation of expected signs in parameter. Bhat (1998) pointed out that the heterogeneous elasticicity in level-of-service (LOS) variables were often ignored in the statistical modeling. The elasticicity in LOS will, in general, vary across sub-groups formed by observed and unobserved individual characteristics. So then in-adequate treatment for the heterogeneity leads into a statistically inferior data fit and also into inappropriate evaluations of policy actions.

Conventional studies indicate that socio-demographic or individual characteristics were very important factors on interregional passenger trip, however the parameters in statistical model sometimes violates the expected condition. Bhatta *et al*. (2011) clarified that LOS variables often include measurement errors, which would cause some biased reproduction on the model of travel demand, due to biased parameter estimation in modal split model. If the critical parameters for policy making do not fulfill the expected condition, it will lead to poor policy decisions. To fix the statistical problem, data sampling methods itself should be revised.

Unfortunately, there are few studies about the improvement of sampling in transportation survey. Yamaguchi (2011) showed that there often occur sampling rate difference between railway and airway in interregional survey in Japan. In this study, they proposed a novel estimation procedure in data expansion coefficient to fit the marginal distribution of individual characteristics with that of population distribution.

1.2 Data Sampling and Propensity Score

The biased parameter estimation occurred in modal split model. However, the most of conventional studies only focuses on the accuracy of the dataset, especially about the transportation service level. On the other hand, the sampling issue for each characteristics is still ignored in most of studies.

The innovation of the randomized experiment was initiated by Ronald Fisher. Randomized experiment can separate the effect of the treated factor on the endogenous variable, from the other factor's effects.

Randomization of samples can be achieved to adjust the covariate distribution to be almost equal between the treated and controlled group. Heckman *et al.* (1995), however, claimed that the randomization in observational surveys about the consumer behavior was not achieved in some case. For example, gender and age are often adjusted as the covariates, while the other socio-demographic characteristics, such as household income or information accessibility are merely adjusted prior to the model estimation, because of the difficulty to get the reference distribution.

Rosenbaum (2002) illustrated two types of bias to give us a wrong implication about the causal effect between the treated variable and endogenous (i.e. interested) variables. One is caused by the difference in observed characteristics in the covariates between treated and controlled group. This bias can be cancelled by making pair of samples with treated and controlled, which have identical (or almost identical) covariates, and then the effect of treatment is calculated by the comparison between two groups. The other one is caused by the difference in non-observed characteristics between the groups. Obviously, the latter bias is not easy to be cancelled.

Rosenbaum *et al.* (1983) proposed a propensity score method to remove the latter bias. Propensity score is the stochastic tendency on the treated variable caused by the set of covariates, which can include the observed and unobserved effect on the covariates. The core idea of propensity score is to adjust the difference in covariate distribution between the treated and controlled group by using the propensity score of each sample. After the first work of Rosenbaum, the propensity score method has been developed and refined in data matching (making pair of data from treated and controlled), data weighting (to adjust the sample weight), and regression analysis.

Bang *et al.* (2005) summarized the effect of propensity score adjustment in causal inference analysis as “doubly robust estimators” about the causation. Doubly robust estimators, give the analyst two opportunities to make a valid causal inference, in contrast to standard statistical analysis. It is well known that in standard regression analysis, the model specification about inclusion / exclusion of the variables would be important to adjust the covariate and treated variable effect on the endogenous variable. If the important covariate is dropped in the model, the causal inference would be easily biased. However in case of introduction of propensity score in regression analysis, the estimator of treated variable would be robust even if the propensity score model is not adequately specified. When the covariates in the propensity score model is not significant, the adjustment effect in regression analysis become random. On the other hand, the covariates is significant, the adjustment by the propensity score will also be effective in the regression model. Bang *et al.* (2005) showed their results by simulation analysis about the finite sample performance of doubly robust estimators, and reported that they have expected performance.

Rubin (1997) discussed the effect of propensity score in large data sets. In his paper, he reported to try to estimate the effects of treatments from such large data sets should be careful in covariate distribution. However, standard methods in statistical softwares (such as linear or logistic regression) can be deceptive for these objectives because they provide no warnings. He argued that propensity score methods were more reliable tools for addressing such objectives.

Propensity score approach can be understood in the missing data treatment in statistical analysis. Table 1 shows the missing data treatment in mortality rate analysis. In this case, treatment variable is “smoking” and endogenous variable is mortality rate increased by smoking.

Table 1. Missing data in the observational survey

mortality rate	Smoking	Non-smoking
Smoker’s	existing	missing
Non-smoker’s	missing	existing

Naïve causal effect can be estimated by the difference of mortality between the treatment group(z=1) and controlled group(z=0). Suppose the mortality of treatment and controlled group referred by y_1 and y_0, respectively. Expected average of treatment effect T will be obtained by eq.(1).

$$T = E(y_1) - E(y_0) \tag{1}$$

In terms of rigorous covariate management between treatment group and controlled group, the ideal comparison requires the counterfactual mortality data in case of smoker’s with non-smoking, or in case of non-smoker’s with smoking. Such comparison gives us the unbiased influence of treatment (i.e. smoking effect on mortality) without the effects from other covariate. Of course such the comparison is impossible. On the other hand, the effect of covariate such as gender, age, occupation, and the daily stress level seem to be significant on mortality, therefore such the indirect effect (from other covariate to mortality via smoking custom) should be removed. Propensity score about the tendency of smoking explained by other covariate can be used for the adjustment.

On non-randomized sampling like the above situation, it is impossible to directly compare the treatment group with controlled group due to the systematic difference in covariates. In order to overcome the limitation of naïve comparison, balancing score b(x) defined as a conditional probability for smoking (treatment condition referred by z) by the other covariate vector x can be utilized. b(x) works to balance the treatment variable z with covariate vector x, in case to fulfill the following orthogonal condition in eq.(2).

$$x \coprod z \mid b(x) \tag{2}$$

Propensity score is conditional probability indicating tendency to assign the experimental treatment. It is called the coarsest balancing score because no other (hidden or multiple) experimental conditions are not explicitly considered but indirectly considered as the unobserved factor. Propensity score is defined as eq. (3).

$$e(x) = p(z = 1 \mid x) \tag{3}$$

Hoshino(2009) introduced several methods how to apply the propensity score analysis in the observational survey data. Propensity score enables the hypothetical sampling under the different treatment. Because true value of propensity score cannot be obtained, it is necessary to estimate from the available data set. The estimation of causal effect by using propensity score is made up by following two steps. Firstly, we should estimate the propensity score model. And then the estimated propensity score is used to adjust the indirect effect of other covariates.

About the second step, Hoshino proposed three ways of bias adjustment procedures following Rosenbaum and Rubin as: matching, stratification, and regression analysis of covariance with IPW index. Among these procedures, regression with IPW index can clarify not only the unbiased treatment effect on the endogenous variable, but also unbiased covariate effect on it. IPW estimator is a weighted average using inverse of propensity score. Let propensity score for sample i e_i, unbiased estimator of y_1 and y_0 will be calculated by eq. (4) and eq. (5), respectively.

$$\hat{E}(y_1) = \sum_{i=1}^{N} \frac{z_i y_i}{e_i} / \sum_{i=1}^{N} \frac{z_i}{e_i} \tag{4}$$

$$\hat{E}(y_0) = \sum_{i=1}^{N} \frac{(1-z_i) y_i}{1-e_i} / \sum_{i=1}^{N} \frac{(1-z_i)}{1-e_i} \tag{5}$$

1.3 Purpose of Our Study

This study purposes to improve the aggregated demand forecasting model of interregional passenger by using propensity score analysis. For this purpose, we apply the propensity score to net passenger traffic survey in Japan. The current problems in demand forecasting model with propensity score are as follows.

1) LOS variables in interregional transportation modes are often correlated between the different modes, due to the spatial characteristics in passenger demand.
2) Observations in passenger demand is not evenly distributed over the LOS distribution
3) Due to the multicolinearity among LOS variables, the estimated parameters of the demand forecasting model do not fulfill the expected sign or significance level.

Table 2. Interregional modal choice and LOS of ODs

mode	Short distance / Higher LOS	Long distance / Lower LOS
Railway (Shin-Kansen)	many	few
Airway	few	many

The observation characteristics in interregional passenger demand in Japan can be summarized as table 2. Comparing with table 1, we can find that the observed passenger demand about the modal choice which is explained by some LOS variables indicates the similar observation distribution with the causation analysis between the smoking and mortality case. Therefore, we can expect that the propensity score analysis can be effectively applied to improve the demand forecasting model of interregional passengers.

2. MODEL SPECIFICATION

Following the Hoshino's procedure, we can set up the propensity score about the composite LOS of each OD. First, we apply principal component analysis for LOS variables, such as railway time, railway distance, railway cost, and railway frequency. Then we calculate principal component score for each sample. Based on the first principal component, which is expected to indicate the composite LOS positively correlated with the substantial distance of the OD. We classify samples into two groups.

$$z_i = \begin{cases} 1, & u_i^1 \geq \tilde{M} \\ 0, & u_i^1 < \tilde{M} \end{cases} \tag{6}$$

$$e_i = P(z_i = 1) = \frac{1}{1 + \exp(-V_i)} \tag{7}$$

where, u_i^1 indicates the first principal component score for sample i. $\tilde{M}$ indicates the median of the score. In eq.(7), V_i indicates the linear propensity function to indicate the tendency for z_i =1. As the explanatory variable we use access time to airport, income (under three million yen, over seven million yen, and unknown), sex, age, trip days, and constant. The parameters of the propensity score can be calculated by logit model.

By using propensity score in eq.(7), IPW index can be calculated in eq.(8).

$$w_i = N \cdot \frac{\dfrac{z_i}{e_i} + \dfrac{1-z_i}{1-e_i}}{\sum_{i=1}^{N} \left(\dfrac{z_i}{e_i} + \dfrac{1-z_i}{1-e_i}\right)} \tag{8}$$

Eq. (8) is the standard IPW weight proposed by Rosenbaum and Rubin. On the other hand in our study, the treatment variable is modal choice itself. Therefore, we try to compare another IPW index to introduce the modal choice variable. The revised IPW can be obtained in eq. (10).

$$y_i = \begin{cases} 1, \operatorname{mod} e : airway \\ 0, \operatorname{mod} e : railway \end{cases} \tag{9}$$

$$w_i^y = N \cdot \frac{\dfrac{y_i}{e_i} + \dfrac{1-y_i}{1-e_i}}{\sum_{i=1}^{N} (\dfrac{y_i}{e_i} + \dfrac{1-y_i}{1-e_i})} \tag{10}$$

where, y_i indicates the airway choice, N indicates sample size and w_i^y indicates the modal-IPW index.

The log-likelihood function of modal choice model can be formulated in eq. (11).

$$\text{logL=}\sum_{\text{i=1}}^{\text{N}} w_i \cdot v_i \left\{ y_i \log(P_i) + (1-y_i)\log(1-P_i) \right\} \tag{11}$$

$$v_i = N \cdot \frac{q_i}{\sum_{i=1}^{N} q_i} \tag{12}$$

where, q_i indicates expansion factor of sample i and v_i indicates the normalized factor by N. By using maximum-likelihood estimation, we estimate the binominal modal choice model formulated in eq. (13). V_i^a is a utility function of airway, while V_i^r is utility function of railway.

$$P_i = P(y_i = 1) = \frac{\exp(V_i^a)}{\exp(V_i^a) + \exp(V_i^r)} \tag{13}$$

$$\hat{A}_j = \sum_{i \in j} P_i \cdot q_i \tag{14}$$

$$\hat{R}_j = \sum_{i \in j} (1-P_i) \cdot q_i \tag{15}$$

$$A_j = \sum_{i \in j} q_i{}^a \tag{16}$$

$$R_j = \sum_{i \in j} q_i{}^r \tag{17}$$

where, the expected passenger trip generation of area j for each mode can be calculated by eq.(14) and eq.(15), respectively. $\hat{A}_j$ is the expected passenger trips by air from area j and to area j, and $\hat{R}_j$ is the expected passenger trips by train from area j and to area j. q_i^a is an expansion factor of i who used air, and q_i^r is an expansion factor of i who used train. $\hat{A}_j$ and A_j ,and $\hat{R}_j$ and R_j are compared to check the reproducibility of the estimated model.

3. DATA

The interregional net passenger traffic survey of Japan started in 1990, and repeated in every 5 years. This survey covers car, railway, airline, and ship passengers from and to all the regions in Japan. The objective of this survey is to provide the fundamental information to plan the interregional transportation infrastructure. Samples are collected by the questionnaire survey for each of modes. Each sample is given the expansion coefficient based on the gross passenger traffic at several control sections, and it is aggregated in the interregional OD tables by each of trip purposes, or each of transportation modes. The aggregated OD tables are available on the website of Ministry of Land, Infrastructure, Transportation, and Tourism of Japan. The summary and report about the survey are also available. After the third survey in 2000, the disaggregated samples with expansion coefficient for each record are available. Interregional tourism passenger traffic data used in this study is rail and air passengers surveyed on a weekday and a holiday in 2005. We use the disaggregated modal choice data with expansion coefficient for each sample. The observation is aggregated for 207 areas covering whole Japan. For simplicity, we extract the passenger traffic data from Tokyo to the other local areas, and from the other local areas to Tokyo. Among the 207 areas, some of isolate islands where do not choose the railway are excluded, so then the 194 areas remain as the sample OD. In order to simplify our analysis, we pick up the ODs from Tokyo to the other areas (Tokyo to Locals), and the other areas to Tokyo (Locals to Tokyo), out of the whole OD pairs.

Since the LOS data of each trip is not surveyed in the above, (due to simplicity and to decrease the respondent’s load), LOS data for each trip is calculated from the shortest path search algorithm. The interregional network in domestic Japan is made from the time table in 2005, which is made up by 525 links and 243 nodes covering the primal railway stations and all the airports.

4. RESULT OF THE MODEL ESTIMATION AND THE PERFORMANCE

4.1 Propensity Score Model

In order to get sample propensity about the LOS (Level-Of-Service) in each OD, we firstly estimated the principal component model to integrate LOS related variables. Table 3 shows the factor loading in first principal component for each variable. The variables used in factor loading are related to railway services which are expected to geographical characteristics sample set. As shown in table 3, signs of factor loading fulfill the expected conditions, because the larger score in first principal component indicates the remoteness of the OD pair, while the smaller score in first principal component indicates the nearness of the OD pair.

Figure 1 and 2 shows the distribution of the first principal component score of each

Table 3. Factor loading in first principal component

Variables	Factor loading	
	Tokyo to Locals	Locals to Tokyo

Railway time	0.976	0.974
Railway distance	0.964	0.963
Railway fare	0.944	0.943
Railway frequency	-0.563	-0.584
Contribution to overall variance	77.27%	77.69%

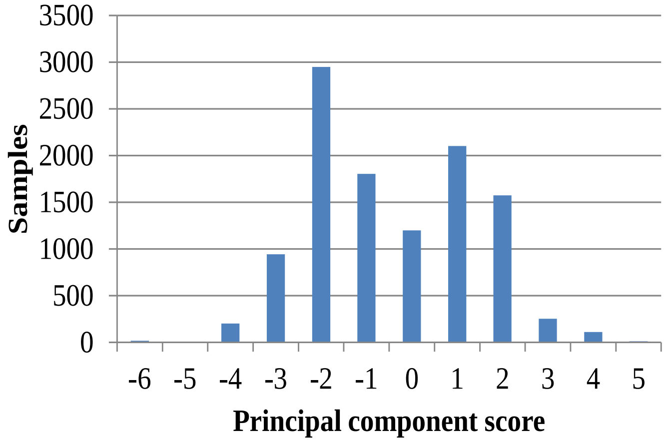

Figure 1. Distribution of the first principal component score (Tokyo to Locals)

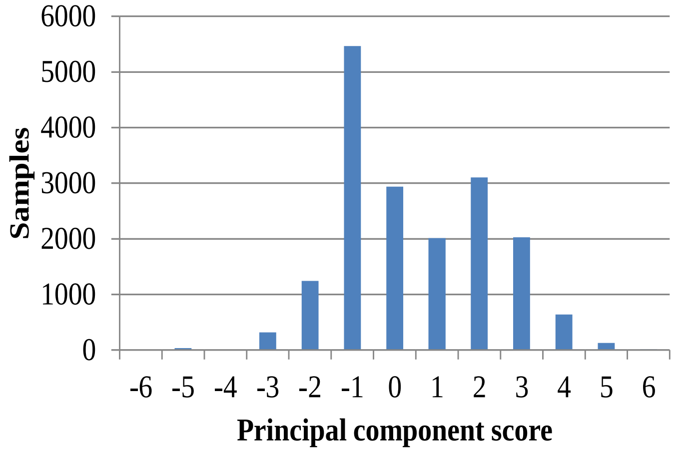

Figure 2. Distribution of the first principal component score (Locals to Tokyo)

sample from Tokyo to Locals, and Locals to Tokyo, respectively. As shown in these figures, the composite LOS of OD pairs do not evenly or normally distribute in the sample set.

Following the procedure in eq.(6), we set the thresholds around the median to divide the sample group into remoteness / nearness in OD characteristics, by setting the treatment variable z. By using z, we estimate the propensity score model for the first principal component in LOS. Table 4 shows the parameters of propensity score model by logit model. As the explanatory variables, we set the covariates relating to the interregional composite LOS. Except the lower income range, the explanatory variables are significant. The model

Table 4. Propensity score model for the first principal component in LOS

	Tokyo to Locals			Locals to Tokyo		
Explanatory variables	Coefficient	t-value		Coefficient	t-value	
Access time to airport	0.001	24.34	***	0.003	48.88	***
Income(under three million yen=1, other=0)	0.136	1.27		-0.045	-0.61	
Income(over seven million yen=1, other=0)	-0.011	-0.22		-0.304	-7.24	***
Income is unknown	-0.159	-1.92		-0.525	-6.97	***
Sex(male=1,female=0)	-0.221	-4.95	***	-0.196	-4.95	***
Age(under 30 years old=1,over 30 years old=0)	0.218	4.74	***	-0.073	-1.80	
Trip days(over two days=1,one day=0)	0.784	18.57	***	0.779	22.01	***
Constant	-1.442	-19.02	***	-2.590	-36.13	***
Initial log-likelihood	-7743.8			-12428.1		
Converged log-likelihood	-7144.8			-9900.9		
Likelihood ratio	0.077			0.203		
Sample size	11172			17930		

: significant at the 5% level; **: significant at the 1% level; *: significant at the 0.1% level*

Table 5. Sample weight and models

Sample weight	Abbreviation
None	NW
IPW by z	WZ
IPW by transportation mode	WM

fitting in log likelihood ratios for both directions are not so low. Based on the estimated probability for composite LOS, we can calculate the weight of each sample from eq.(8) and eq.(10). Table 5 shows the abbreviations to refer the each weighting method.

Figure 3 and 4 shows the propensity weight with expansion coefficient for each weighting aggregated for short and long distance, in case of Tokyo to Locals. In order to classify each sample depending on its remoteness, we used the airway distance as a reference to set the thresholds as 700 km from origin to destination. As shown in these figures, the weight calculated from eq.(8) will give the significant change in sample weight comparing with the non-weighted case. And the weight calculated from eq.(10) will give a moderate modification to the non-weighted case. Such the characteristics are also seen in the Locals to Tokyo case.

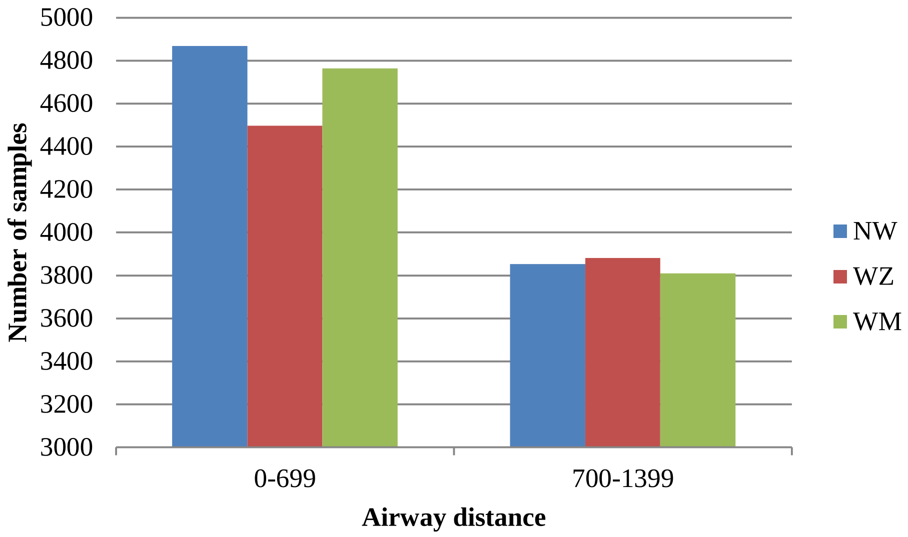

Figure 3. Weight for railway (Tokyo to Locals)

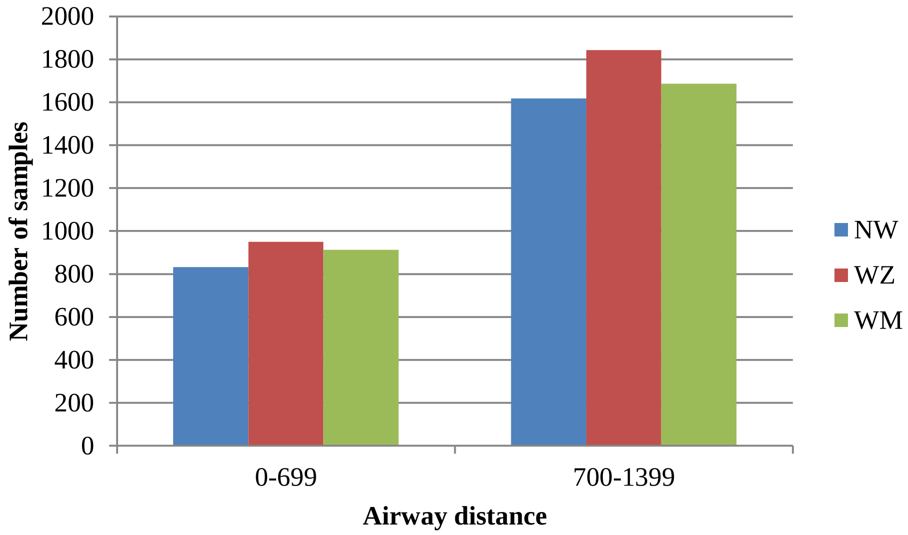

Figure 4. Weight for airway (Tokyo to Locals)

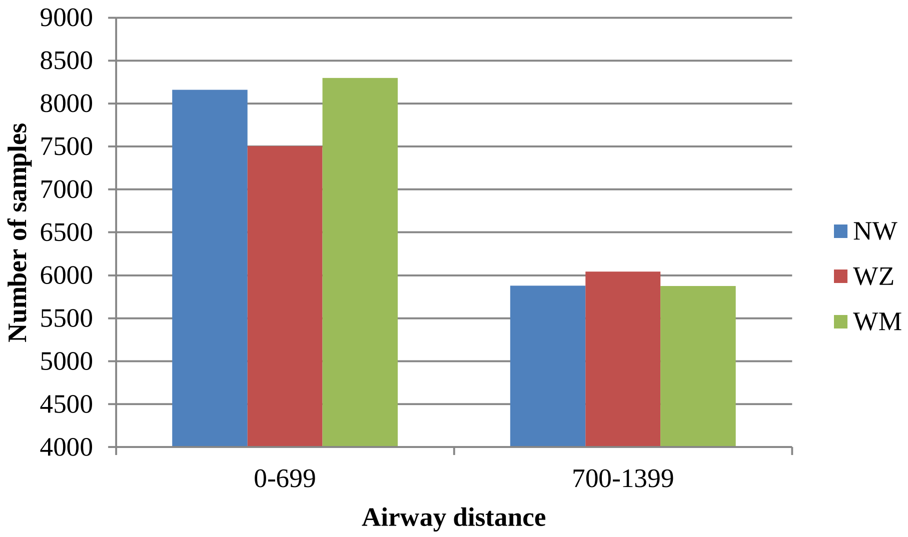

Figure 5. Weight for railway (Locals to Tokyo)

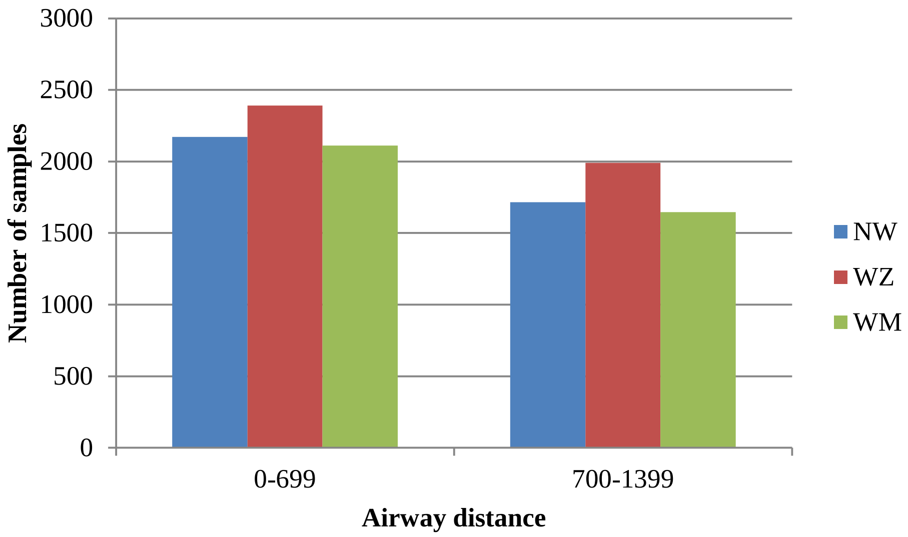

Figure 6. Weight for airway (Locals to Tokyo)

4.2 Modal Split Model and Reproductivity of Data

Table 6 and 7 show the estimated parameters of modal split model by using eq. (8). For comparison, the non-weighted model is also shown in the tables. Compared with the non-weighted case, by introducing propensity score, log-likelihood ratios for both direction (Tokyo to Locals, and Locals to Tokyo) are decreased. The expected signs of LOS parameters are negative in time and distance for both modes, while they are positive in the frequency for both modes. About the non-weighted model in Tokyo to Locals, airway time, airway distance and airway frequency do not fulfill the above conditions. In the non-weighted model in Locals to Tokyo, railway frequency, airway distance and airway frequency do not fulfill the above condition. By introducing the propensity score and weighted by treated condition z (i.e. composite LOS), the sign of parameters in table 6 are identical between NW and PS-WZ models. On the other hand in table 7, the sign of parameters are improved in both frequencies, but worsened in airway time.

Table 6. Estimated parameters in modal split model (Tokyo to Locals, weighted by z)

	NW			PS-WZ		
Explanatory variables	Coefficient	t-value		Coefficient	t-value	
Railway time	-3.077	-7.74	***	-2.107	-6.85	***
Railway distance	-0.005	-9.38	***	-0.006	-14.70	***
Railway frequency	0.003	0.30		0.014	1.79	
Airway time	0.108	0.43		1.088	5.41	***
Airway distance	0.001	3.43	***	0.000	0.63	
Airway frequency	-0.012	-2.16	*	-0.017	-3.72	***
Income(under three million yen=1, other=0)	-0.520	-1.92		-0.524	-2.22	*
Income(over seven million yen=1, other=0)	0.764	6.84	***	0.805	8.30	***
Income is unknown	-0.482	-2.57	*	-0.483	-2.83	**
Sex(male=1,female=0)	-1.005	-9.99	***	-1.047	-11.81	***
Age(under 30=1,other=0)	0.124	1.29		0.202	2.50	*
Constant	-21.506	-11.02	***	-21.257	-13.30	***
Initial log-likelihood	-7743.8			-7743.8		
Converged log-likelihood	-2218.0			-2342.3		
Log of Likelihood ratio	0.714			0.697		
Sample size	11172			11172		

: significant at the 5% level; **: significant at the 1% level; *: significant at the 0.1% level*

Table 7. Estimated parameters in modal split model (Locals to Tokyo, weighted by z)

	NW			PS-WZ		
Explanatory variables	Coefficient	t-value		Coefficient	t-value	
Railway time	-2.982	-10.11	***	-2.690	-10.36	***
Railway distance	-0.004	-11.16	***	-0.005	-14.73	***
Railway frequency	-0.008	-1.09		0.000	0.01	
Airway time	-0.169	-0.82		0.473	2.90	**
Airway distance	0.001	5.17	***	0.000	2.88	**
Airway frequency	-0.008	-1.77		0.003	0.71	
Income(under three million yen=1, other=0)	-1.043	-6.52	***	-1.156	-9.74	***
Income(over seven million yen=1, other=0)	1.054	12.92	***	1.035	14.90	***
Income is unknown	-1.279	-9.92	***	-1.260	-9.46	***
Sex(male=1,female=0)	-1.152	-15.58	***	-1.119	-17.51	***
Age(under 30=1,other=0)	0.314	4.20	***	0.310	4.78	***
Constant	-19.785	-12.39	***	-21.273	-14.57	***
Initial log-likelihood	-12428.1			-12428.1		
Converged log-likelihood	-4302.6			-4321.2		
Log of Likelihood ratio	0.654			0.652		
Sample size	17930			17930		

: significant at the 5% level; **: significant at the 1% level; *: significant at the 0.1% level*

Table 8. Estimated parameters in modal split model (Tokyo to Locals, weighted by mode)

	NW			PS-WM		
Explanatory variables	Coefficient	t-value		Coefficient	t-value	
Railway time	-3.077	-7.74	***	-3.217	-9.67	***
Railway distance	-0.005	-9.38	***	-0.004	-10.33	***
Railway frequency	0.003	0.30		0.005	0.52	
Airway time	0.108	0.43		-0.116	-0.54	
Airway distance	0.001	3.43	***	0.000	0.63	
Airway frequency	-0.012	-2.16	*	-0.008	-1.68	
Income(under three million yen=1, other=0)	-0.520	-1.92		-0.669	-2.94	**
Income(over seven million yen=1, other=0)	0.764	6.84	***	0.778	8.41	***
Income is unknown	-0.482	-2.57	*	-0.245	-1.64	
Sex(male=1,female=0)	-1.005	-9.99	***	-0.731	-8.61	***
Age(under 30=1,other=0)	0.124	1.29		-0.057	-0.70	
Constant	-21.506	-11.02	***	-20.125	-12.23	***
Initial log-likelihood	-7743.8			-7743.8		
Converged log-likelihood	-2218.0			-2551.7		
Log of Likelihood ratio	0.714			0.670		
Sample size	11172			11172		

: significant at the 5% level; **: significant at the 1% level; *: significant at the 0.1% level*

Table 9. Estimated parameters in modal split model (Locals to Tokyo, weighted by mode)

	NW			PS-WM		
Explanatory variables	Coefficient	t-value		Coefficient	t-value	
Railway time	-2.982	-10.11	***	-3.656	-15.38	***
Railway distance	-0.004	-11.16	***	-0.003	-9.04	***
Railway frequency	-0.008	-1.09		-0.007	-1.01	
Airway time	-0.169	-0.82		-1.071	-6.62	***
Airway distance	0.001	5.17	***	0.000	-0.45	
Airway frequency	-0.008	-1.77		0.003	0.77	
Income(under three million yen=1, other=0)	-1.043	-6.52	***	-0.982	-7.90	***
Income(over seven million yen=1, other=0)	1.054	12.92	***	1.415	22.23	***
Income is unknown	-1.279	-9.92	***	0.548	6.18	***
Sex(male=1,female=0)	-1.152	-15.58	***	-0.906	-15.16	***
Age(under 30=1,other=0)	0.314	4.20	***	0.406	7.17	***
Constant	-19.785	-12.39	***	-16.904	-13.46	***
Initial log-likelihood	-12428.1			-12428.1		
Converged log-likelihood	-4302.6			-5210.8		
Log of Likelihood ratio	0.654			0.581		
Sample size	17930			17930		

: significant at the 5% level; **: significant at the 1% level; *: significant at the 0.1% level*

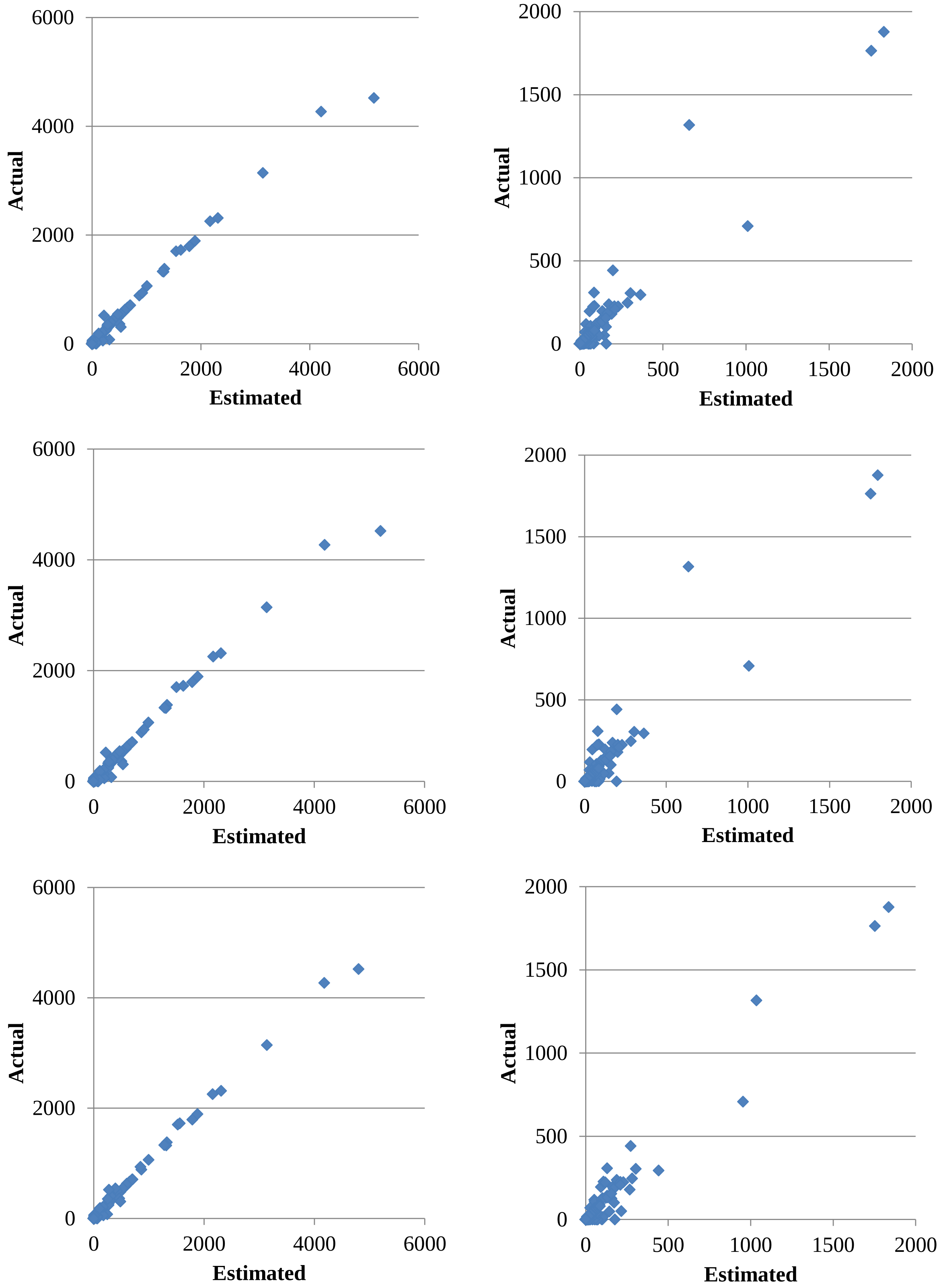

Figure 7. Reproducibility:Tokyo to Locals (left:rail,right:air/top:NW,middle:WZ,bottom:WM)

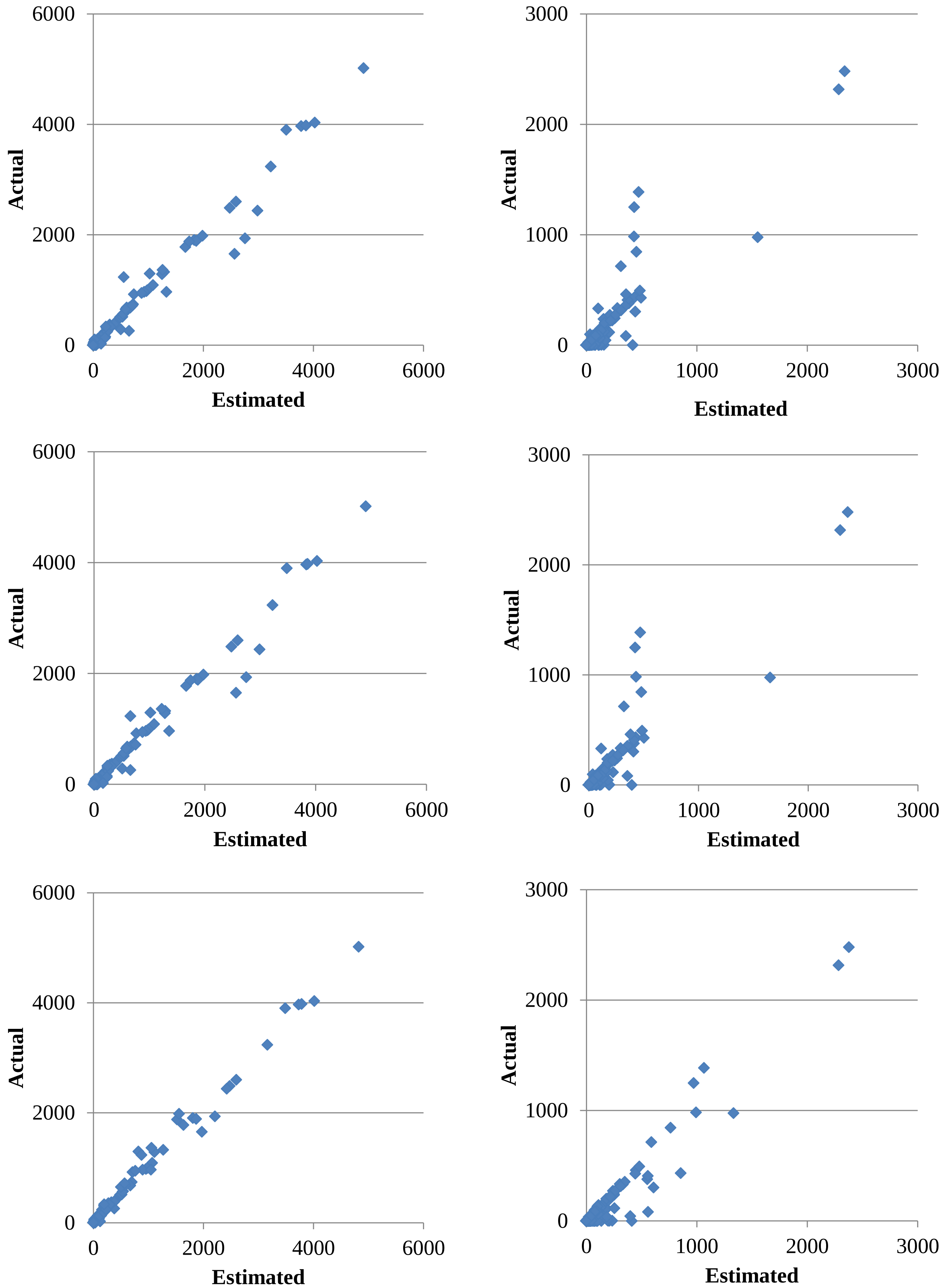

Figure 8. Reproducibility:Locals to Tokyo (left:rail,right:air/top:NW,middle:WZ,bottom:WM)

Table 8 and 9 show the estimated parameters of propensity score model by using eq. (10). For comparison, the non-weighted model is shown again in both tables. By introducing propensity score, log-likelihood ratios for both direction (Tokyo to Locals, and Locals to Tokyo) are decreased. The expected sign of LOS parameters are identical to table 6 and 7. In table 8, the sign of parameters are improved in airway time. In terms of airway distance and airway frequency, the sign of parameters still violate the expected condition, but the significance level is decreased. In table 9, the sign of railway frequency still violates the expected condition, while the airway time (to become significant), airway distance (to become insignificant with violating expected condition), and airway frequency (the sign of parameters become improved) are improved by introducing propensity score calculated by eq.(10).

In order to confirm the parameter collection effect by propensity score weight by mode, figure 8 and figure 9 show the reproductivity of each model aggregated in 194 areas, Tokyo to Locals, and Locals to Tokyo, respectively. In both figures, left side shows the railway reproductivity, while right side shows the airway reproductivity. And also in both figures, the top is calculated from non-weight model, the middle is calculated from the propensity score in eq.(8) weighted by composite LOS (z), and the bottom is calculated from the propensity score in eq.(10) weighted by mode. As shown in both figures, the demand reproductivity of PS-WM model is significantly better than other models. Comparing the PS-WZ model and the PS-WM model, the former reproductivity is quite similar with non-weighted case for both modes, while the latter is much improved, i.e. the scatter plot between the observed and estimated data close to linear relation along with the diagonal line.

The reason why such the difference in reproductivities among the models can be considered as follows: in the WZ model, the propensity of LOS is directly connected with composite weight of LOS distribution by eq.(8) , but in the WM model, the propensity of LOS is connected with modal choice by eq.(10). Since we estimated the modal split model, responsiveness to LOS is not affected by composite LOS but by the modal share itself. Therefore, the sample weight modification by eq.(10) gives a better performance than that by eq.(8), and the reproductivity is improved, even the index of model fitness (log-likelihood ratio) is decreased.

5. CONCLUSION

This study purposes to improve the aggregated demand forecasting model of interregional passenger by using propensity score analysis. For this purpose, we apply the propensity score to net passenger traffic survey in Japan. The current problems in demand forecasting model with propensity score are as follows.

The empirical analysis of the proposed method clarified that the simple application of propensity method does not give the significant improvement in the reproductivity of aggregated area, while the propensity score weighting performs well, in case that the inverse of propensity weight is made for modal choice variable. In the latter case, the expected conditions of modal choice model become to be improved, and also the reproductivity of the sample is improved.

The remaining issue is to check the applicability of the proposed method for other dataset, as to include other OD pairs. Another issue is to find the more effective treated variable other than composite railway LOS, and to check the condition that such the improvement occurs, by referring to the covariate distribution characteristics.

REFERENCES

Bhat, C.R. (1998) Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. *Transportation Research*, 32(7), 495-507

Bhatta, B.P., Larsen, O.I. (2011) Errors in variables in multinomial choice modeling: A simulation study applied to a multinomial logit model of travel mode choice. *Transport Policy*, 18(2), 326-335

Fisher, R.A. (1935) The design of experiments. Edinburgh, UK: Oliver & Boyd.

Fu, X., Oum, T.H., Jia, Y. (2012) An Analysis of Travel Demand in Japan's Inter-city Market: Empirical Estimation and Policy Simulation. *Journal of Transport Economics and Policy*, 48, 1-28.

Gone, X., Yan, H., Liu, S., Xu, Z. (2011) Tourism traffic mode choice behavior model. *ICCTP 2011*, 766-775.

Heckman, J., Smith, J. (1995) Assessing the case for social experiments. *Journal of Economic Perspectives*, 9, 85-110.

Hoshino, T. (2009) Statistical science for observational data. Iwanami-shoten, (in Japanese).

Kaneko, Y., Kato, H. Review of studies for inter-urban transportation in Japan.

Kato, H., Endo, K., Kurita Z., Kaneko, Y., Kato, K., Shimizu, T., Tanabe, K. (2011) Inter-regional travel demand analysis using integrated model. *Journal of the Eastern Asia Society for Transportation Studies*, 9, 239-254.

Okumura, M., Nakagawa, M., Yamaguchi, K., Tsuchiya, K., Okumura, Y., Hino, S., Makoto, T. (2002) Topics on analysis and evaluation of inter-urban transportation. *Proceedings of Infrastructure Planning*, 25.

Rosenbaum, P.R. (2002) Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, 17, 286-304.

Rosenbaum, P.R., Rubin, D.B. (1983) The central roles of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41-55.

Rubin, D.B. (1997) Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, 127, 757-763.

Shenyang, G. Fraser, M.W. (2010) Propensity score analysis. SAGE Publication,Inc.

Yamaguchi, H., Okumura, M., Okubo, K. Study on a Methodology to Set the Magnification Factor for Inter-city Travel Samples. *Proceedings of Infrastructure Planning*, 43.