Journal of the Eastern Asia Society for Transportation Studies, Vol. 6, pp. 173 - 185, 2005

TRANSFERABILIY AND UPDATING ANALYSIS OF MODE CHOICE MODEL BETWEEN TWO DEVELOPING COUNTRIES – PREDICTIVE PERFORMANCE POINT OF VIEW

Djoen San SANTOSO
Graduate Student
Department of Civil Engineering
Saitama University
255 Shimo-Okubo, Sakura-ku
Saitama-shi, Saitama
338-8570 Japan
Fax: + 81-48-858-3555
E-mail: sgd2055@post.saitama-u.ac.jp

Koji TSUNOKAWA
Professor
Department of Civil Engineering
Saitama University
255 Shimo-Okubo, Sakura-ku
Saitama-shi, Saitama
338-8570 Japan
Fax: + 81-48-858-3825
E-mail: koji-t@post.saitama-u.ac.jp

Abstract: This paper presents the predictive performance analysis of transferring and updating a mode choice model estimated in Ho Chi Minh City to Phnom Penh. Difficulties in transferring a model due to different data collection method and simplified approaches to solve the problems are also presented. Four updating approaches associated with small sized sample of 200, 400 and 600 observations are applied in this study. The study shows that predictive performance improves with updating alternative specific constants and improves further with additional adjustment of scale parameter. Combined transfer estimator has proved to produce the most significant improvement in updating the model. In case of the existence of transfer bias, application of Bayesian updating is not recommended.

Key Words: transferability, updating approach, mode choice model, Ho Chi Minh City, Phnom Penh.

1. INTRODUCTION

Transportation sector has an important role in supporting economic developments of a country. Transportation has been a connector and a link of two different areas, both for transporting people and goods. A well-planned and systematic transportation system will improve the efficiency of the whole system. Many developing countries are in the stage of building their transportation infrastructures and systems to support the economic development. However, financial constraints and availability of resources have been the major concerns in this sector.

An alternative approach in the development of transportation modeling would be in the line with the effort to help developing countries to improve their transportation sectors. Transferability of mode choice model can be considered as one of the alternatives. Many studies have tried to analyze and explore this subject, mostly using data from developed countries (Atherton and Ben-Akiva, 1976; Talvitie and Kirshner, 1978; Koppelman and Wilmot, 1982; Galbraith and Hensher, 1982; Koppelman, F.S. *et al.*, 1985; Gunn, H.F. *et al.*, 1985; Ben-akiva and Bolduc, 1987; Badoe and Miller, 1995a and 1995b; Karasmaa, 2001). In 2005, Santoso and Tsunokawa examined the transferability of mode choice model using data from a developing country, travel survey data of Ho Chi Minh City was used in their case study. Statistical criteria were applied to analyze the results. The study showed that naïvely transferring a model is not recommended. Updating approaches need to be applied in transferring the model to improve the performance of the model, and, at least, 400 observations of small sample data should be used for updating purposes.

However, transferring a model to a new area is not as simple as expected. Differences in transportation characteristics between the area where the model was estimated (estimation context) and the area where the model is going to be transferred (new application context) influence the transferability process in term of the judgment to include or exclude variable(s) of the model in the new context (Santoso, 2004; Santoso and Tsunokawa, 2005). Another obstacle is the different method applied in collecting survey data. There is no standardization in conducting data collection, such as in sampling procedure, definition of variables, measurement procedure and content of questionnaire, which likely have influenced the estimated model coefficients (Galbraith and Hensher, 1982). This problem has also influenced the availability of travel survey data needed for updating purpose.

From the above problem statements, an examination of the application of transferability in actual situation, in the scope of developing countries, is necessary, which is chosen as the main topic of this study. This study explores the empirical analysis of transferability and updating approaches between two developing countries, which is the extension of previous study conducted by Santoso and Tsunokawa (2005). Mode choice model estimated from the travel survey data of the urban area of Ho Chi Minh City (HCMC), Vietnam, is transferred to Phnom Penh City in Cambodia. The primary objective of the study is to systematically compare the performances of updating methods in transferring a mode choice model between two developing countries assessed by their predictive abilities. It is also expected, as the secondary objective, that the study can illustrate difficulties in transferring a mode choice model between two countries.

The next section of this paper briefly explains the methodology of the study. The data used in this study and comparison of the two cities are provided in section 3. In section 4, the transferability and updating methods are discussed, and then section 5 describes the data and results of the analysis. The final section summarizes the findings of the study and provides recommendations for further study as well.

2. RESEARCH METHODOLOGY

This study focuses on journey to work trip data. A multinomial logit model for mode choice estimated in the urban area of HCMC from the previous study (Santoso and Tsunokawa, 2005) is used as a model to be transferred to the new context, in this study, Phnom Penh City in Cambodia. The model is referred to as HCMC model. In HCMC model, three modes are considered: walking, bicycles and motorcycles. These three modes were the first three dominant modes for the journey to work. The other modes, including automobile and public bus, each has an insignificant share. Therefore they were not included in the analysis. As we are going to transfer the HCMC model to Phnom Penh, these three modes are the focus of our analysis. Phnom Penh is chosen not without reasons. This city has quite similar mode share as HCMC with motorcycle as the most dominant mode for the journey to work. Walking also has a significant share in this trip purpose. Although, there is only a small portion of people riding bicycle for going to work, it would be a challenge to find out how this situation may affect the result of transferability. Both cities are also characterized with high mixed traffic, as well as low availability and quality of public transportation.

In building the HCMC model, many variables generated from travel survey data have been tried to be incorporated by considering the expected signs of coefficients and the statistical values. Finally, eight variables are used as the explanatory variables of the HCMC model, which are presented in Table 1. The model was estimated from a total of 10,353 observation data, which covered 16,740 households.

Table 1 Variables of HCMC Mode Choice Model

Variable	Definition
BASC	1 for bicycle mode 0 otherwise
MASC	1 for motorcycle mode 0 otherwise
TRAVT	travel time (minute)
COST/INC	travel cost (vietnamese dong) / worker's monthly income (vietnamese dong) for motorcycle mode 0 otherwise
MOWN	1 if worker owns a motorcycle for motorcycle mode 0 otherwise
BOWN	1 if worker owns a bicycle for bicycle mode 0 otherwise
MALE	1 for male worker for motorcycle mode 0 otherwise
VEHRAT	number of vehicles available in the household / number of workers in the household for walk mode 0 otherwise

Transferability study in a developing country carried out by Santoso and Tsunokawa (2005), which based on statistical criteria, shows that naïvely transferring a model is not recommended; updating approaches should be integrated in the transferability process to get a better model. They also proposed that a small sample not less than 400 observations should be used for updating purpose to reduce possible negative effects of large variance as a result of small sample size. Based on their results, this study transfers the HCMC model to the Phnom Penh City with four updating procedures: updating alternative specific constants (ASCs), updating ASCs and scale parameter, Bayesian updating, and combined transfer estimator; at the same time, three sets of small sample sizes of 200, 400 and 600 observations generated randomly from Phnom Penh travel survey data are integrated in the analysis to verify the result of their study. This study focuses on the analysis of predictive performance in judging the usefulness of the models instead of statistical criteria.

However, due to the different data collection method applied in Phnom Penh, some data are not available or not exactly available to update the variables of HCMC model. Given this situation, simplified alternative procedures were developed to solve the problems, which are explained as follows:

1. Travel time. The available information in Phnom Penh data is only the travel time of the selected mode by traveler. To solve this problem, the average travel speeds of the three

modes are used to convert the available travel time of the selected mode to the other two modes.

2. Travel cost (for motorcycle). The travel cost data are not completely available for all motorcycle users. Only 3% of motorcycle travelers provided valid information about their travel costs. In dealing with this problem, a linear regression relationship between the available travel cost and travel time data is developed. The relationship is used to estimate the unavailable travel cost data.
3. Number of workers in the household. In Phnom Penh survey data, there is no information about the number of workers in each household. The available information, which is quite close to this variable, is the number of residents and the number of residents over 5 years old. As the number of workers in the household together with the number of vehicles variables are used to represent the availability of vehicles for the household members to accommodate their travel needs, the information about the number of residents over 5 years old is likely to provide the same representative to achieve the purpose.
4. Male. There is no information about the gender of travelers and an alternative approach seems not possible to solve this problem so this variable is excluded in the transferability process. As this study tries to analyze transferability and updating as practical as possible, we just take the gender variable out of the model and keep the coefficients of other variables as they are. There is no re-estimation of the model. One thing should be noted here, as we remove one variable from the model without re-estimating the model, bias may occur in the model, which may have either positive or negative effect to the application context. Ideally, we should re-estimate the model or another approach is to replace the parameter with the average value. However, as both options require the availability of data of the original context, which may not be practical in real world, neither of them is apply in this study.

The above approaches are incorporated in the transferability process. Performances of the updated models are examined by their predictive ability. Each model is applied to the whole to-work trip data to estimate the mode share. The estimated mode share (prediction) is compared with the observed modal split (observation) to judge the predictive performance of the model. By this way, we can examine which updating method and small sample size produce good results.

3. HO CHI MINH AND PHNOM PENH CITIES

Phnom Penh is the capital of the Kingdom of Cambodia, a country with area of 181,035 km^2. In 2002, the GDP per capita of the country was US$ 297, which was around 60% of Vietnam GDP per capita in 2003 (US$ 485). The travel survey data of Phnom Penh was part of the Study of Transport Master Plan on the Phnom Penh Metropolitan Area conducted by JICA, which commenced in 2000; meanwhile, the HCMC model was estimated from the travel survey data conducted in 2002.

The similarity of the two cities, in terms of mode choice, can be seen from the role of motorcycle as the dominant mode in daily commuting activities. In addition, in both cities, we observe that many people travel to their destinations on foot, which is the second highest selected mode. However, while bicycles have a considerable share in traveling activities in HCMC, it is not the case in Phnom Penh. Bicycle share in Phnom Penh is very low, which contributed only 1.9% of the total mode share for all travel activities.

Comparison of characteristics of the study areas between HCMC and Phnom Penh is presented in Table 2.

4. TRANSFERABILITY AND UPDATING APPROACHES

Many researches have suggested that updating approaches need to be applied to improve the performance of the model to be transfer (transfer model). Naïvely transferring a model is not a preferable option because no model is ever sufficiently specified (Galbraith and Hensher, 1982), which also means no model can perfectly reflect the travel behavior of commuters. As a result, assessing model transferability only on the basis of the set of model parameters being equal in the two areas is unlikely to be met (Ben-Akiva, 1981; and Koppelman and Wilmot, 1986).

Table 2 Characteristics of HCMC and Phnom Penh

	HCMC (Urban)	**Phnom Penh**
Location	Southeast Asia	Southeast Asia
Population	3,611,329	1,101,918*
Area (km2)	142	374
Number of surveyed households	16,740	6,455
Average HH member	4.1	5.8
Number of observations	26,864	4,956
GDP per capita (US$)	1,640**	215
Mode share (%): Motorcycle	84.17	77.6
Walking	7.99	22.1
Bicycle	7.84	0.3
Daily trips per person	3.0	2.0
Increase rate of population (annual)	2.1	2.7

Note: all data based on the year the surveys were conducted except stated on (HCMC: 2002, Phnom Penh: 2000)
* 1998 data
** the value was for all HCMC, not limited to the urban area.

Updating approaches are used to modify parameters of the transfer model by incorporating available information about the new application context. Thus, in applying updating methods, small sample data from the new application context should be available. Many updating approaches have been developed to help improving the performance of the model to be transfer. The most prominent one is by updating the alternative specific constants. In addition to this updating method, adjusting the scale parameter could also be included in the process to improve further the model. As have been identified by McFadden (1976): and Westin and Manski (1979), there are three types of differences that may exist in models between estimation and application contexts: differences in the alternative specific constants, in the sensitivity or scale of the model parameters, and in the relative values of variable coefficients.

Another approach in updating method is by combining the coefficients of the model from the estimation context with those estimated on the small sample data of the new application context. There are two updating methods which apply this approach in their calculation: Bayesian updating and combined transfer estimator. In 1976, Atherton and Ben-Akiva proposed to use Bayesian theorem in estimating new parameters of the application context. Parameters from the model estimated using the available small sample are combined with parameters of the transfer model to yield asymptotically normal updated parameters. The formula to calculate the updated coefficients and standard deviations are as follows:

$$\beta_u = \left(\Sigma_o^{-1} + \Sigma_s^{-1}\right)^{-1} \left(\Sigma_o^{-1}\beta_o + \Sigma_s^{-1}\beta_s\right) \tag{1}$$

and

$$\Sigma_u = \left(\Sigma_o^{-1} + \Sigma_s^{-1}\right)^{-1} \tag{2}$$

where

β_u = [K x 1] vector of updated parameters, where K = M + N -1, where
M is the number of explanatory variables and
N is the number of alternatives
β_o = parameter vectors of the transferred model
β_s = parameter vectors derived from the small sample
Σ_u = covariance matrix of updated parameters
Σ_o = covariance matrix of parameters of the transferred model
Σ_s = covariance matrix of parameters estimated from the small sample

The method was used with considerable success in their research study.

Bayesian updating implicitly assumed that the underlying set of parameters are expected to be equal: $E[\beta_o] = E[\beta_s] = E[\beta_u]$. This assumption may be difficult to be justified in case of transferring a model to a very different context including transferring over a long period of time. To overcome this problem, Ben-Akiva and Bolduc (1987) developed a combined transfer estimator which is based on the mean squares error criterion and extends the Bayesian procedure to explicitly account for the presence of a transfer bias. The approach accounts for a non zero transfer bias and yields the minimum mean square error estimate of the updated parameters achievable from a linear combination of the estimation and application context parameter estimates. The minimum mean square error estimate is provided by:

$$\beta_u = \left(\left(\Sigma_o + \Delta\Delta^T\right)^{-1} + \Sigma_s^{-1}\right)^{-1} \left(\left(\Sigma_o + \Delta\Delta^T\right)^{-1}\beta_o + \Sigma_s^{-1}\beta_s\right) \tag{3}$$

where
$\Delta = \beta_o - \beta_s$ (know as transfer bias)
Δ^T = transpose of Δ

The approach, given the transfer bias is small, is expected to yield better estimates of updated parameters than simply using the parameters of the model estimated from small sample as applied in Bayesian updating.

5. DATA ANALYSIS AND RESULTS

The four updating methods discussed above: updating ASCs, updating ASCs and scale parameter, Bayesian updating and combined transfer estimator are used in the transferability process of this study. Their performances in improving the transfer model are examined to find which updating approach provides better improvement. Santoso and Tsunokawa (2005), in their transferability study in Ho Chi Minh City, suggested that small sample data not less than 400 observations should be used to reduce the possible negative effects of the large small sample variance. This study works in this line and tries to investigate further the validation of the result. Totally, there are 4,956 to-work trip observation data generated from the travel survey of Phnom Penh. Besides the minimum suggested small sample size of 400 observations, observation data of 200 and 600 are also used to investigate the performance of the number of observations less and more than the suggested minimum number of observations. We use three sample sets for each sample size, each of which is generated randomly from the travel survey data of Phnom Penh. The structure of small samples and updating approaches used in this study is presented in Figure 3. Additionally, it is reasonable to assume that as the size of small sample increases, the performance or accuracy of the updated models also improves. Based on this assumption, the whole to-work trip observation data is also employed in the updating processes, which the results are used as references.

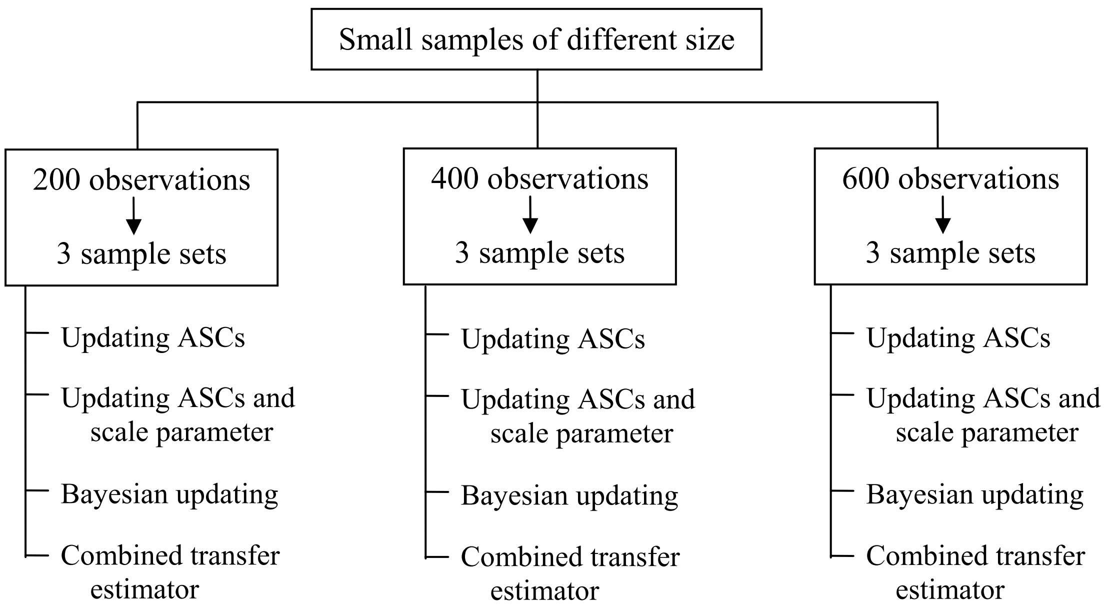

Figure 3 Structure of Small Samples and Updating Approaches Used in the Study

In this study, instead of based on statistical criteria to judge the model, evaluation of the model performance is conducted by examining the predictive ability of the model. The evaluation of the results is conducted by comparing the predicted modal split of the model with the observation. Each model, which is updated using small sample data of Phnom Penh, is evaluated using the whole sample data (4,956 observations) to compute the mode share, and then the result is compared with the observed modal split of the entire data. The predicted modal share is estimated by calculating individual probabilities of choosing the available modes and then summing these probabilities over the whole sample.

The results of predictions from the models are presented in Table 3. The rows of the table show the predicted mode share for each mode per set of small sample with the last three rows are the results of updating using full sample data of Phnom Penh. The applied updating methods are presented starting from the fifth column. The prediction of mode share by naïvely transferring the model is also included in the table (the fourth column). The purpose is, using performance of the naïve method as a benchmark, to evaluate the performance of alternative updating approaches. As stated before, the observed mode share for walking, bicycle and motorcycle are 1095 (22.1%), 15 (0.3%) and 3846 (77.6%), respectively. To provide clear description of the results and to show performance of the models, relative errors between prediction and observation are calculated, as shown in Table 4. Relative error is computed as the absolute difference between prediction and observation divided by observation. The overall weighted average of the relative error for each small sample size, which is computed by summing the weighted relative error of each set of small sample divided by number of sets (three for this study), is presented in Figure 4. The observed shares of the modes are used as the weight in the computation. The formula is presented as follows:

$$OWA_z = \sum_{i=1}^{3} \{(O_w \, x \, RE^i_{wz}) + (O_b \, x \, RE^i_{bz}) + (O_m \, x \, RE^i_{mz})\} / 3 \qquad (4)$$

where:

OWA_z = overall weighted average of relative error for Z observation
O_w = observed walk share
O_b = observed bicycle share
O_m = observed motorcycle share
RE^i_{wz} = relative error of walk for the *i*th small sample set of Z observation
RE^i_{bz} = relative error of bicycle for the *i*th small sample set of Z observation
RE^i_{mz} = relative error of motorcycle for the *i*th small sample set of Z observation

Figure 4 at the small sized sample of 200 observations with updating method of combined transfer estimator shows intriguing result. In detail, we can refer to Tables 3 and 4, in the first set of 200 small samples and under the combined transfer estimator method, we found that only for this set of small sample, combined transfer estimator did not perform well. This may be due to the quality of small random sample generated was not good enough to represent the whole sample, which strongly correlated with the size of small sample. The low share of bicycles should also be counted in adding problem of generating random sample to sufficiently represent this mode. With this situation, the existence of outlier should be noted in evaluating the performances of updating methods, particularly for sample size in question.

In general, as expected, the predictive performance of the naïve method did not produce good result. Figure 4 shows that performance of Bayesian updating is not satisfactory. This finding supports the result of Santoso and Tsunokawa (2005). It is observed that the disappointing result of Bayesian updating is due to the existence of transfer bias between the coefficients of the transfer model and the model estimated from small sample data, as suggested by Gunn, H.F. *et al.* (1985) that Bayesian updating should be used when the transfer bias is assumed to be negligible. Surprisingly, the relative errors of Bayesian updating are quite similar to the naïve method (columns 4 and 7 of Table 3). As we know, small sample data has a large variance as its negative characteristic. This is also the case with variances of coefficients of models updated from small samples of 200, 400 and 600 observations (values are not presented in this paper). The large variance has made the updated coefficients from Bayesian updating closer to coefficients of transfer model, which make the coefficients of the model from small sample have no significant contribution in updating the model. This reason explains why Bayesian updating generated comparable proportion of modal split to the naïve method.

Table 3 The Predicted Mode Share

Small sample	Set	Mode	Predicted Mode Share									
			Naïve		ASCs		ASCs & Scale		Bayesian		Combined	
				(%)		(%)		(%)		(%)		(%)
200	1	Walking	600	12.11	1340	27.04	1128	22.76	602	12.15	428	8.64
		Bicycle	471	9.50	71	1.43	405	8.17	475	9.58	3974	80.19
		Motorcycle	3885	78.39	3545	71.53	3423	69.07	3879	78.27	554	11.18
	2	Walking	600	12.11	1158	23.37	939	18.95	601	12.13	1039	20.96
		Bicycle	471	9.50	104	2.10	260	5.25	472	9.52	11	0.22
		Motorcycle	3885	78.39	3694	74.54	3757	75.81	3883	78.35	3906	78.81
	3	Walking	600	12.11	1217	24.56	983	19.83	601	12.13	964	19.45
		Bicycle	471	9.50	119	2.40	226	4.56	470	9.48	63	1.27
		Motorcycle	3885	78.39	3620	73.04	3747	75.61	3885	78.39	3929	79.28
400	1	Walking	600	12.11	1273	25.69	985	19.87	606	12.23	946	19.09
		Bicycle	471	9.50	82	1.65	144	2.91	475	9.58	162	3.27
		Motorcycle	3885	78.39	3601	72.66	3827	77.22	3875	78.19	3848	77.64
	2	Walking	600	12.11	1240	25.02	1028	20.74	606	12.23	1066	21.51
		Bicycle	471	9.50	64	1.29	234	4.72	471	9.50	31	0.63
		Motorcycle	3885	78.39	3652	73.69	3694	74.54	3879	78.27	3859	77.87
	3	Walking	600	12.11	1395	28.15	1046	21.11	609	12.29	1028	20.74
		Bicycle	471	9.50	46	0.93	107	2.16	468	9.44	62	1.25
		Motorcycle	3885	78.39	3515	70.92	3803	76.74	3879	78.27	3866	78.01
600	1	Walking	600	12.11	1298	26.19	993	20.04	610	12.31	1021	20.60
		Bicycle	471	9.50	53	1.07	203	4.10	476	9.60	46	0.93
		Motorcycle	3885	78.39	3605	72.74	3760	75.87	3870	78.09	3889	78.47
	2	Walking	600	12.11	1234	24.90	1019	20.56	613	12.37	1082	21.83
		Bicycle	471	9.50	55	1.11	244	4.92	475	9.58	32	0.65
		Motorcycle	3885	78.39	3667	73.99	3693	74.52	3868	78.05	3842	77.52
	3	Walking	600	12.11	1277	25.77	1000	20.18	615	12.41	1061	21.41
		Bicycle	471	9.50	67	1.35	78	1.57	473	9.54	52	1.05
		Motorcycle	3885	78.39	3612	72.88	3878	78.25	3868	78.05	3843	77.54
Full		Walking	600	12.11	1273	25.69	1035	20.88	705	14.23	1082	21.83
		Bicycle	471	9.50	23	0.46	76	1.53	506	10.21	15	0.30
		Motorcycle	3885	78.39	3660	73.85	3845	77.58	3745	75.56	3859	77.87

Table 4 Relative Error between Prediction and Observation

Small Sample	Set	Mode	Relative Error between Prediction and Observation				
			Naïve	ASCs	ASCs & scale	Bayesian	Combined
200	1	Walking	0.452	0.224	0.030	0.450	0.609
		Bicycle	30.400	3.733	26.000	30.667	263.933
		Motorcycle	0.010	0.078	0.110	0.009	0.856
	2	Walking	0.452	0.058	0.142	0.451	0.051
		Bicycle	30.400	5.933	16.333	30.467	0.267
		Motorcycle	0.010	0.040	0.023	0.010	0.016
	3	Walking	0.452	0.111	0.102	0.451	0.120
		Bicycle	30.400	6.933	14.067	30.333	3.200
		Motorcycle	0.010	0.059	0.026	0.010	0.022
400	1	Walking	0.452	0.163	0.100	0.447	0.136
		Bicycle	30.400	4.467	8.600	30.667	9.800
		Motorcycle	0.010	0.064	0.005	0.008	0.001
	2	Walking	0.452	0.132	0.061	0.447	0.026
		Bicycle	30.400	3.267	14.600	30.400	1.067
		Motorcycle	0.010	0.050	0.040	0.009	0.003
	3	Walking	0.452	0.274	0.045	0.444	0.061
		Bicycle	30.400	2.067	6.133	30.200	3.133
		Motorcycle	0.010	0.086	0.011	0.009	0.005
600	1	Walking	0.452	0.185	0.093	0.443	0.068
		Bicycle	30.400	2.533	12.533	30.733	2.067
		Motorcycle	0.010	0.063	0.022	0.006	0.011
	2	Walking	0.452	0.127	0.069	0.440	0.012
		Bicycle	30.400	2.667	15.267	30.667	1.133
		Motorcycle	0.010	0.047	0.040	0.006	0.001
	3	Walking	0.452	0.166	0.087	0.438	0.031
		Bicycle	30.400	3.467	4.200	30.533	2.467
		Motorcycle	0.010	0.061	0.008	0.006	0.001
Full		Walking	0.452	0.163	0.055	0.356	0.012
		Bicycle	30.400	0.533	4.067	32.733	0.000
		Motorcycle	0.010	0.048	0.000	0.026	0.003

The results also indicates significant improvement obtained by updating alternative specific constants, which strengthen the statement of Koppelman, F.S. *et al.* (1985) that adjustment of alternative specific constants appears to be a universally desirable approach. Additionally, if we further improve the model by also adjusting the scale parameter, excluding the result of small sample of 200 observations, better improvement can be achieved but smaller in magnitude, as shown in Figure 4.

The combined transfer estimator, with the exception of the result of small sample of 200 observations, produced the most satisfactory results. The inclusion of transfer bias in the calculation procedure of this method has important role in producing this achievement. The method produced slightly better improvement than updating ASCs and scale parameter.

This study, despite the fact that outlier is present in small sized sample of 200 observations, could not confidently make any conclusion related to the appropriate minimum small sample for updating. However, it is observed that for updating ASCs, updating ASCs and scale parameter and combined transfer estimator, the tendency shows that as the size of the small sample increases, the prediction accuracy also increases, which at the end will reach to the value produced by updating using full sample data.

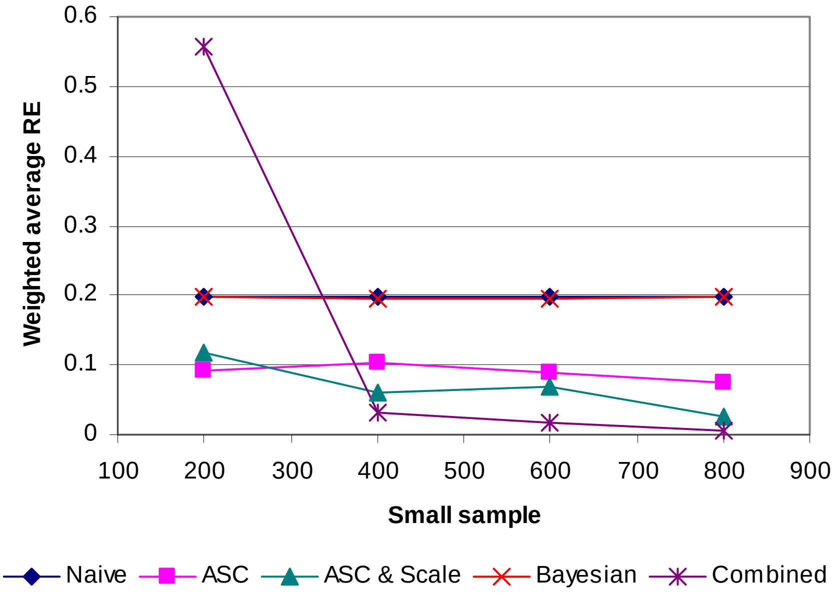

Figure 4 Small Sample, Updating Method and Relative Error

6. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER STUDY

This study empirically examines the spatial transferability and updating analysis of two developing countries, Ho Chi Minh City (HCMC) and Phnom Penh. Multinomial logit model for mode choice estimated from the journey to work trip data of the urban area of HCMC was transferred to Phnom Penh. Four updating approaches: updating alternative specific constants, updating alternative specific constants and scale parameter, Bayesian updating and combined transfer estimator, were applied in the transferability procedure together with small-sized sample of 200, 400 and 600 observations. To decrease bias in the result, three sets of random sample were generated for each sample.

Some problems exist related to the availability of data from the travel survey in Phnom Penh to update the variables of HCMC model. Due to different data collection method applied in Phnom Penh, data of number of residents over 5 years old was used as a proxy for the number of workers. Travel time data are available only for the chosen mode, so we estimated the travel times of unchosen modes by multiplying the travel time of chosen mode with the ratios of average travel speeds of the modes. In addition, travel cost data for motorcycle users are rarely available. The relationship between the available travel cost of chosen mode and travel time of the mode in question was used to generate a linear regression formula, which, then, was applied to estimate the missing travel cost data. One variable of HCMC model, the gender variable, had to be excluded before transferring the model because there is no information about the gender of travelers in Phnom Penh data and alternative approach was not applicable to be developed to solve the problem.

The results of the present paper could not provide confirmation that small sample data not less than 400 observations should be used for updating purposes as yielded by Santoso and

Tsunokawa (2005). However, it is concluded that the predictive performances of the updating methods improve with the number of observations.

Updating alternative specific constants significantly improve the model predictability and the predictive performance improves further with additional adjustment of scale parameter. Due to the existence of transfer bias, Bayesian updating did not produce satisfactory results. On the other hand, combined transfer estimator, as the extension of Bayesian updating, produced considerably good prediction of mode share, owing to the inclusion of transfer bias in its estimation.

To further transferability study, analysis of transferability in terms of statistical criteria for this context of study should also be conducted. It would be interesting to compare results from the statistical criteria and results from this study to gain insight whether both points of view produce similar conclusions or not. Applications of this study in other cities besides Phnom Penh would also be interesting to be observed as another suggestion for further study. It is hoped that the present study can contribute in stimulating the interest for advancing transferability study, particularly in the context of developing countries.

ACKNOWLEDGEMENTS

The authors are thankful to JICA and ALMEC Corporation for permitting the use of the travel survey data of HOUTRANS and Phnom Penh for this research study.

REFERENCES

Atherton, T.J., and Ben-Akiva, M.E. (1976) Transferability and Updating of Disaggregate Travel Demand Models, **Transportation Research Record No. 610**, 12-18.

Badoe, D.A., and Miller, E.J. (1995a) Analysis of the Temporal Transferability of Disaggregate Work Trip Mode Choice Mode, **Transportation Research Record No. 1493**, 1-11.

Badoe, D.A., and Miller, E.J. (1995b) Comparison of Alternative Methods for Updating Disaggregate Logit Mode Choice Models, **Transportation Research Record No. 1493**, 90-100.

Ben-Akiva, M. (1981) Issues in Transferring and Updating Travel-Behavior Models. In **New Horizons in Travel-Behaviour Research** (Stopther, P.R., Meyburg, A.H., and Brog, W., eds.), D.C. Heath and Co., Lexington, Massachusetts.

Ben-Akiva, M., and Bolduc, D. (1987) Approaches to Model Transferability and Updating: The Combined Transfer Estimator, **Transportation Research Record No. 1139**, 1-7.

Galbraith, R.A., and Hensher, D.A. (1982) Intra-metropolitan Transferability of Mode Choice Models, **Journal of Transport Economics and Policy, Vol. 16, No. 1**, 7-29.

Gunn, H.F., Ben-Akiva, M.E., and Bradley, M.A. (1985) Tests of the Scaling Approach to Transferring Disaggregate Travel Demand Model, **Transportation Research Record No. 1037**, 21-30.

Karasmaa, N. (2001) The Spatial Transferability of the Helsinki Metropolitan Area Mode Choice Models. **Proceedings of the 9th WCTR**, Seoul, 2001.

Koppelman, F.S., and Wilmot, C.G. (1982) Transferability Analysis of Disaggregate Choice Models, **Transportation Research Record No. 895**, 18-24.

Koppelman, F.S., Kuah, G.K., and Wilmot, C.G. (1985) Transfer Model Updating with Disaggregate Data, **Transportation Research Record No. 1037**, 102-107.

Koppelman, F.S., and Wilmot, C.G. (1986) The Effect of Omission of Variables on Choice Model Transferability. **Transportation Research Part B, Vol. 20, No. 3**, 205-213.

McFadden, D. (1976) Properties of the Multinomial Model. Working Paper 7617. Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.

Santoso, D.S. (2004) Mode Choice Model of Developing Countries – Characteristics and Transferability. **Proceedings 10th ISSOT Symposium**, Hanoi, Vietnam, 16-18 November 2004.

Santoso, D.S., and Tsunokawa, K. (2005) Transferability and Updating of Mode Choice Model in Developing Countries – Case Study of Ho Chi Minh City. **The 84th Transportation Research Board Annual Meeting**, Washington, D.C., 9-13 January 2005.

Talvitie, A., and Kirshner, D. (1978) Specification, Transferability and the Effect of Data Outliers in Modeling the Choice of Mode in Urban Travel, **Transportation, Vol. 7**, 311-331.

Westin, R., and Manski, C.F. (1979) Theoretical and Conceptual Developments in Demand Modelling. In **Behavioral Travel Modelling** (Hensher, D.A., and Stopher, P.R., eds.), Croom-Helm, London.