A STUDY ON MODAL CHOICE CHARACTERISTICS BASED ON RP AND SP APPROACH IN METRO MANILA

Hisao UCHIYAMA
Associate Professor
Department of Civil Engineering
Science University of Tokyo
2641 Yamazaki Noda-city,
Chiba Pref. Japan
Fax: +81-471-23-9766

Yuichi MOHRI
Research Associate
Department of Civil Engineering
Science University of Tokyo
2641 Yamazaki Noda-city,
Chiba Pref. Japan
Fax: +81-471-23-9766

abstract: This study focuses on the applicability of the disaggregate approach with small sampling size to developing countries. The study is (1) to analyze the mode choice characteristics respectively using RP and SP data in Metro Manila, (2) to illustrate the difference in the choice among the segments, then (3) to refer to the homogeneity of the choice after calibrating a combined model simultaneously using RP and SP data. The study concludes to use not single data set from RP or SP but the combined data set, because the assumption of "homogeneity" may tend to overcome that of "heterogeneity".

1. INTRODUCTION

A metropolitan area in most developing countries such as Metro Manila has faced on serious socio-economic troubles caused by the traffic congestion with an increase in car usage. In order to relieve this current chaos of the congestion, the area necessitates new introduction of railway system, for instance, Light Rail Transit (LRT) from a long-term viewpoint, as well as implementation of measures regarding Transportation System Management (TSM) or Transportation Demand Management (TDM) from a short-term viewpoint. In fact, Metro Manila already operated LRT #1 in 1985, and intends to open line #2 and #3 in near future.

For the establishment of appropriate urban transportation planning, it is very important primarily to grasp the change in mode choice behavior brought by the presence of newly added service to a mode. This study tries to develop the method for an analysis of the mode choice characteristics in these urban areas. It is usually difficult, however, to analyze the characteristics on the basis of collecting a large quantity of sampled data, because of the severe budget constraint of these countries. A disaggregate approach is said, on the other hand, to offer much useful information about the behavioral characteristics even in small sampling size in data collection.

A macroscopic analysis based on this approach applied to much wider spread urban area in space requires that the characteristics obtained from relatively less samples must reflect that from the population. This means that the "homogeneity" about behavior (choice), which corresponds to the response to the attributes such as fare, travel time and so on, must be guaranteed. Otherwise if the behavior should be assumed to be "heterogeneous" by segment of the samples, sampling size will increase in accordance with the number of the segments.

Uchiyama, H. et al., introduced an exogenous function with a couple of variables of fare and trip distance into a modal split model in Metro Manila using RP (revealed preference) data, then suggested that different segments in trip distance resulted to be homogeneous in their behavior. It seems, however, to be unreasonable to discriminate between the "homogeneity" and the "heterogeneity" only from the relation between the revealed choice obtained from the RP data and the variables of LOS (level of service such as cost, in-vehicle-time, etc.) given exogenously.

An analysis of the modal behavior using SP (sated preference) data, recently developed, may have a good chance to explain directly the relation between perception of preference and service attributes of each alternative mode, since the SP model is calibrated on the basis of direct response to various hypothetical service attributes of alternatives. Therefore the use of the SP data has an advantage in discussing about them with much more reasonableness than that of the RP data only.

This study focuses on the applicability of the disaggregate approach with small sampling size to developing countries, in which the study is (1) to analyze the mode choice characteristics using RP and SP data in Metro Manila as a case study, (2) to illustrate the heterogeneity of the choice among the segments, which will be derived from the RP data, then (3) to refer to the homogeneity of the choice after calibrating SP models. A combined model simultaneously using RP and SP data, which was proposed by Morikawa, T et al., is applied for further discuss about the homogeneity.

2. CONDUCT OF QUESTIONNAIRE SURVEY

2.1 Study Area

Metro Manila, which consists of 7 cities and 10 municipalities, is currently divided into 202 traffic zones for transportation and city planning purposes. The questionnaire survey is conducted in 67 zones out of the above whole area, where are located not only along the trajectory of the existing LRT #1 but also in the immediate vicinity of the proposed LRT line #2 as shown in Figure 1. This Figure likewise illustrates the boundaries of the cities and the municipalities, the major roads, the train lines of the Philippine National Railway (PNR), the 15-kilometer long LRT line #1 in the direction of north and south, and the 10-kilometer long LRT line #2 in the direction of east and west. The target sample size is, 1,500 households per street. Enumerators, who were asked to visit the randomly picked households and interview all household members, were hired to conduct the survey proper from April 20 to May 8, 1992.

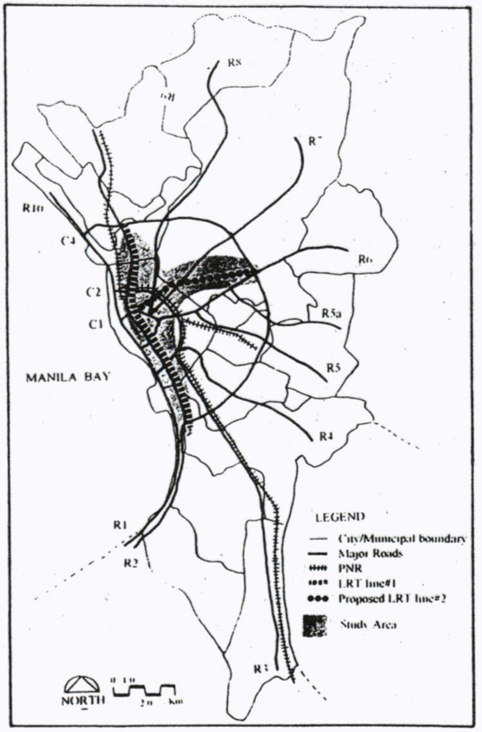

Figure 1 Coverage of the Survey Area

2.2 Questionnaire Form

Basically three types of questionnaire forms are prepared in the survey. the first form contains questions regarding the socio-economic attributes of the household at one level, and the socio-economic attributes of each household member on another.

The second, which is related to the RP, asks about the regular trip making patterns of all household members who at the time of the interview hold regular jobs. The questions asked during this phase of the interview are as follows: regular departure time from home, usual arrival time at the place of work, type of mode regularly used, e.g.., jeepney, bus, LRT, or private vehicle (hereinafter referred as "car"), transferring points (if transfers are made), and frequency of use. The work trip makers are also asked about regular alternative routes and modes used and the primary reason for using them. Answers to questions regarding alternative modes are later referred to in the making the alternative sets and generation on the LOS data.

The third, which is related to the SP, asks about the preferable mode to the proposed LRT line #" when it would be operated. The question are composed of 3 respective binary choice (LRT-Car, LRT-Jeepney, and LRT-Bus) under the condition of 3 different segments of travel time distance to the work place (short-, medium-, and long-trip). Then the respondent compares 9 combinations of hypothetical levels of services (fare, travel time, walking time, and waiting time) provided by the LRT with those of the other existing modes, and answers his/her preference out of following 5 categories: definitely prefer LRT, probably prefer LRT, no preference, probably prefer anther mode, definitely prefer anther mode. Although the respondent results to be requested to give the answer by 3 mode choice, by 3 travel distance and by 9 LOS combination, he/she is permitted to reply them only regarding his/her actual travel distance which was already known at the time of the RP survey. Hence 81 (3x3x9) different stated preference cards are prepared, and 27 cards with the same travel distance category out of 81 are displayed to a respondent in order to help his/her smooth derivation of the answer from complicated comparisons. Figure 2 shows an example of the SP card.

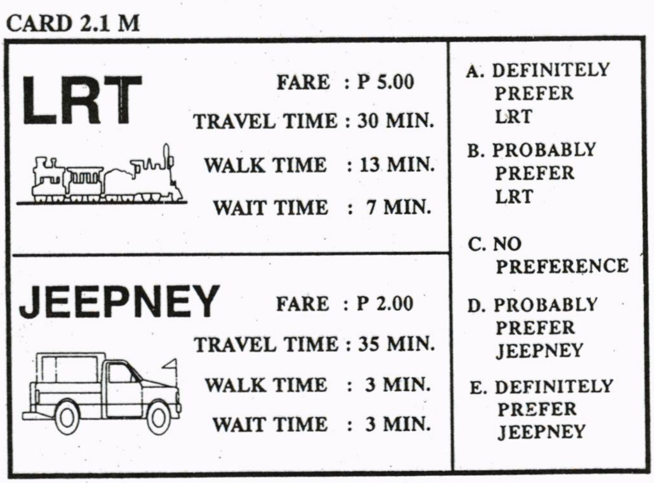

Figure 2 Exampleof SP Card

2.3 Overall Profile of Data Collected

The total number of households interviewed was 1,513, 13 more than the target household sample size. The total number of respondents who registered work trips was 1,234, where

749 came from the area #1 that is along LRT line #1, 762 from area #2, while from the overlapped area 277 were obtained who were already counted in the above a couple of areas as shown in Table 1. Out of this number, 232 walked to work or used some other minor transit mode like motorbike, tricycle, etc. The total number of respondents who either used a major transit mode, i.e., jeepney, bus, LRT, or a car was 1,002, which is equal to the number of samples for the RP.

On the other hand, Table 2 shows the total number of respondents for the 3 respective modal choices of the SP survey. Maximum expected number of respondents, therefore, would reach to that of the RP survey to be equal to 1,234 for respective modal choice. In comparison with each number among the three choice sets, the number of LRT-Jeepney choice is larger than those of others due to higher availability of the jeepney. Furthermore, because a respondent has an opportunity to answer his/her preference for 9 different combination of the LOS during the SP survey, the number of responses or samples for the SP would count 9 times of the number in the Table.

Table 1 Profile of Samples Collection

Area	Number of Zone	Number of Household	Number of Respondent	Number of Sample
Area #1	45	823	749	606
Area #2	38	1,028	762	583
Overlapped	16	338	277	187
Overall	67	1,513	1,234	1,002

Table 2 Samples Collected for SP Survey

Choice Set	LRT - Jeepny			LRT - Bus			LRT - Car		
Distance	Short	Medium	Long	Short	Medium	Long	Short	Medium	Long
Area #1	186	257	40	22	55	16	12	32	4
Area #2	115	202	135	9	37	70	13	78	64
Overall	274	385	151	25	84	74	22	96	66
	810			183			184		

3. MODAL CHOICE CHARACTERISTICS BASED ON RP DATA

3.1 Trip Data Applied

The modal share of main modes by area is shown in Figure 3. Though the jeepney usage occupies a much higher percentage in area #1 and #2 than the others, the difference in the share between the LRT usage and that of the car in both areas appears according to high availability of the existing LRT #1. The average travel distance and its distribution by main mode for overall samples is shown in Figure 4. The jeepney tends to be used in relatively shorter distance trip, and the bus and LRT in medium distance trip, while, the car is consistently used in all distance categories.

The fundamental premise of the disaggregate analysis is "the rational choice" which means a trip maker must have at least more than two alternatives, otherwise he/she must be regarded as "captive". Using the answers taken from the questions regarding the availability of alternative modes and the choice sets generated during the LOS data making process, a cross-tabulation between main mode and captivity is generated and shown in Table 3. This table shows that the sample set can be categorized into three groups; (1) those captive by circumstance, where the respondent has no other feasible alternative mode plying between his home and his place of work so is forced to take the mode he is currently using; (2) those captive by choice, where the respondent has some feasible alternative modes but always take only the current mode he is using; and (3) non-captive, where the respondent actually uses some of the alternative modes in given in his choice set.

There are a lot of possible reasons why respondents in the "captive by choice" category exhibit this travel behavior. Perhaps they perceive that the mode they are currently using is a lot safer, or more comfortable, or more convenient compared to the alternative, or perhaps they are totally unaware of existing alternatives. In this sense, samples except for belonging to the above category (1), captive by circumstance, are to be applied to the following RP analysis.

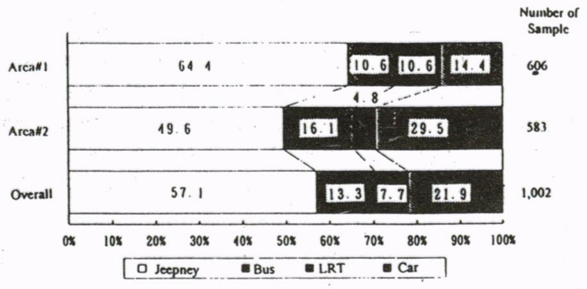

Figure 3 Modal Share of Main Mode

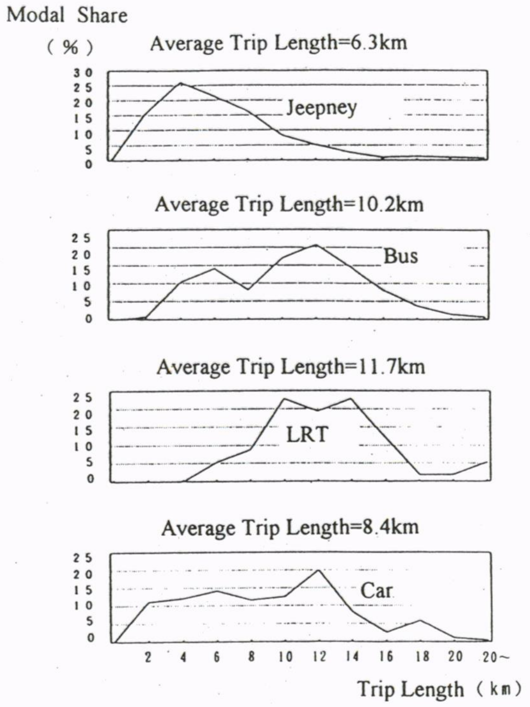

Figure 4 Travel Distance and Distribution

Table 3 Main Mode vs Captive

Mode	(1)*	(2)*	(3)*	TOTAL
Jeepney	53	171	394	573
Bus	6	29	98	133
LRT	0	25	52	77
Car	3	54	162	219
TOTAL	62	279	661	1,002

* Categories:
(1) - captive by circumstance
(2) - captive by own choice
(3) - non-captive

3.2 Modal Choice Models

Several modal choice models for the work trips assumed as a multinominal choice of 4 modes are calibrated by area (area #1 and area #2) and by trip distance (short- (<5km), medium- (5-10km), and long- (>10km) distance category). Estimated parameters and calibration statistics are listed in Table 4. which shows that of goodness-of-fit and Hit-ratio are relatively high. The COST parameter by trip distance displays an increasing tendency from the short- to the medium-distance category, and eventually breaks down in the long-distance category, i.e., changes into positive sign.

Since the COST parameter is not estimated stably in the multinominal choice models, then the binary choice models of LRT and Jeepney are applied. In addition, both modes of LRT and Jeepney are recognized as especially important public transportation modes in Metro Manila from the following viewpoints: (1) new LRT lines are now planned and the SP analysis will be done for the proposed line #2 in the next, and (2) the Jeepney, an indigenous paratransit mode in the Philippines, is very frequently used as shown in Figure 3.1. The result of the calibration is tabulated in Table 5, in which statistics of ρ^2 and Hit-ratio are estimated high similarly to the above multinominal models. However, the COST parameter is also instable and insignificant because of the low t-value for all distance categories and the positive sign for the medium-distance category.

Table 4 Estimated Statistics of Multinominal Choice Models with RP Data

Variable	Area #1			Area #2			Overall model
	Short-dist model	Medium-dist model	Long-dist model	Short-dist model	Medium-dist model	Long-dist model	
COST	-0.152 (-1.29)	-0.059 (-0.71)	0.554 (2.99)	-0.347 (-2.32)	-0.107 (-2.27)	0.043 (0.88)	-0.05 (-3.08)
IVTT	-0.555 (-1.04)	-0.037 (-2.57)	-0.08 (-4.74)	-0.007 (-0.11)	-0.014 (-0.89)	-0.05 (-0.51)	-0.049 (-8.51)
WALK	-0.047 (-0.56)	-0.045 (-0.91)	-0.11 (-1.46)	-0.074 (-0.67)	-0.048 (-1.04)	-0.05 (-0.88)	-0.073 (-3.09)
OVTT	-0.541 (-3.11)	-0.225 (-3.12)	-0.521 (-4.35)	-0.244 (-1.30)	-0.31 (-4.41)	-0.754 (-6.28)	-0.309 (-8.56)
Jeepney CONSTANT	0.904 (0.84)	0.044 (0.003)	18.522 (3.71)	-1.044 (-0.86)	-0.294 (-0.03)	3.223 (2.09)	1.375 (3.77)
Bus CONSTANT	1.508 (0.91)	-0.511 (-0.29)	17.617 (3.69)	-2.528 (-1.32)	0.088 (0.08)	5.447 (3.33)	0.958 (2.21)
LRT CONSTANT			14.411 (3.44)		-1.125 (-1.05)	5.522 (3.35)	-0.644 (-1.46)
Hit Rate	84.1	86.5	79.1	81.5	83.2	85.5	83.6
ρ^2	0.547	0.542	0.641	0.557	0.545	0.685	0.622
$\bar{\rho}^2$	0.519	0.53	0.625	0.527	0.533	0.679	0.62
Sample	88	192	110	81	173	186	812

IVTT : in-vehicle-time . OVTT : out-of-vehicle-time = waiting time + transfer time . WALK : walking time

Table 5 Estimated Statistics of Binary Choice Models with RP Data (LRT-Jeepney)

Variable	Short-dist model	Medium-dist model	Long-dist model	Overall model
COST	-0.5604 (-0.87)	0.1569 (0.88)	-0.1022 (-0.23)	-0.0469 (-0.34)
IVTT	-0.0964 (-3.49)	-0.0866 (-5.71)	-0.0915 (-2.40)	-0.0815 (-7.05)
OVTT	-0.1653 (-2.35)	-0.17 (-2.07)	-0.4043 (-1.88)	-2.023 (-3.04)
CONSTANT	-2.647 (-2.16)	-4.506 (-4.54)	-2.827 (-1.40)	-3.715 (-4.99)
Hit Rate	94.74	91.77	96.3	92.82
ρ^2	0.436	0.401	0.439	0.394
$\overline{\rho}^2$	0.419	0.39	0.384	0.389
Sample	133	231	54	418

IVTT : in-vehicle-time
OVTT : out-of-vehicle-time = waiting time + transfer time + walking time

3.3 Findings from RP Analysis

The result of the multinominal analysis suggests that because the COST parameter is more sensitive to trip distance than the others, there may exist certain threshold of trip length among short-, medium-, and long-distance for mode choice. Such inconsistent estimated statistics in each segment model naturally balances out in the overall model in the right side of Table 4. This balancing effect might tend to hide the true general mode selection behavior of trip makers in Metro Manila. The result of the binary analysis generally implies the same tendency as that of the multinominal analysis except for the statistics of the COST parameter of the overall model. The above findings may direct toward some trials of the further analysis under the assumption of "homogeneity", although the several RP models here were calibrated by segment which means "heterogeneity",

4. MODAL CHOICE CHARACTERISTICS BASED ON SP DATA

4.1 Screening Process of SP Data

A respondent answers only one preferred alternative which is actually used (revealed) in the RP survey. In the SP survey, he/she can give a lot of preferences of which one is obtained from a certain experimental (stated) condition selected among various hypothetical situations. Hence the SP data can convey a larger amount of information on individual preferences than the RP data. The SP data, however, might contain some systematic error caused by misunderstandings of the respondent for example, which is usually called as SP bias. The RP data, therefore, are said to have higher validity than the SP data in this sense. The lack of this validity, which is equivalent to the existence of the SP bias, is known as a disadvantage point on SP approach. Since the bias is principally brought by the quality of the questionnaire or the way of setting the experiments, the bias in this study is to be defined as follows; (1) completely same answering pattern (the answer is always the same, e.g., A. definitely prefer LRT in Figure 2.2 for 9 different preference cards), to which many of Jeepney riders as mode captive belong to, (2) inconsistent answer for changes in various levels of the services.

The above inadequate samples are removed according to the screening process as shown in Table 6. The process in the Table traces the decreased number of adequate samples and the improved F-statistics of each LOS variable step by step. This process is applied to the LRT-Jeepney choice and the F-value is obtained from the Variance Analysis. The number of samples after the Step 1 decreases to about 60-75% for that of overall samples excluding those with lack of answer. Finally the number of samples after Step 2 is left only about 20-25% of all samples. F-statistics of each variable after the screening process is improved to indicate statistically significant at 5% level.

Table 6 Screening Process and F-value for SP Data

(Total samples excluding lack of answer)

	Short-dist	Medium-dist	Long-dist
IVTT	2.98**	3.83**	1.03
WALK	0.52	0.3	0.38
WAIT	1.28	0.65	0.85
COST	0.92	1.97**	0.93
Sample	2,228 (100%)	3,465 (100%)	1,359 (100%)

** indicates significant at 5 % level

⇓

(Step 1) Samples except for answers with completly same pattern

⇓

(Samples after screening of Step 1)

	Short-dist	Medium-dist	Long-dist
IVTT	3.93**	5.45**	1.77
WALK	0.76	0.4	0.6
WAIT	1.73	0.81	1.32
COST	1.29	2.53**	1.68
Sample	1,575 (71%)	2,574 (74%)	846 (62%)

** indicates significant at 5 % level

⇓

(Step 2) Samples except for inconsistent answers with change in LOS

⇓

(Samples after screening of Step 2)

	Short-dist	Medium-dist	Long-dist
IVTT	7.32**	9.7**	4.39**
WALK	2.19**	4.87**	1.89
WAIT	2.33**	3.9**	1.32
COST	2.55**	7.06**	2.35**
Sample	468 (21%)	882 (26%)	279 (21%)

** indicates significant at 5 % level

4.2 Modal Choice Models

The modal choice models are calibrated using the screened SP data aforementioned which are the preferences for the proposed LRT line #2. The SP data prepared here are from a data set of LRT-Jeepney binary choice, because the number of samples of the set (1,629) is the largest compared with those of the other 2 sets (603 of LRT-Bus and 504 of LRT-Car respectively). Furthermore, the SP data actually applied here are those which respond to the alternative of extremely discriminative 2 categories (either "definitely prefer LRT" or "definitely prefer Jeepney") of 5 categories on the questionnaire form, and the data responding to the rest of 3 categories ("probably prefer LRT", "no preference" and "probably prefer Jeepney") are removed, because the response to these 3 categories is scarcely sensitive than that of the former against hypothetical changes in the LOS. Thus the number of samples used decreases from 1,629 to 816. The estimated parameters by trip distance shown in Table 7 tend to be similar to those of the same binary models from the RP analysis as shown in Table 5. This time, however, all the LOS variables have negative coefficients with statistically significant t-value in each distance model and the overall model. Although ρ^2 of the good-of-fit and Hit-ratio depreciate, especially the COST parameters become more stable as indicated by t-statistics for all models.

Table 7 Estimated Statistics of Binary Choice Models with SP Data (LRT-Jeepney)

Variable	Short-dist model	Medium-dist model	Long-dist model	Overall model
COST	-0.8714 (-3.16)	-0.6814 (-3.97)	-1.3724 (-3.15)	-0.8262 (-6.03)
IVTT	-0.1357 (-2.49)	-0.1089 (-3.27)	-0.1678 (-2.41)	-0.1203 (-4.61)
OVTT	-0.2772 (-2.07)	-0.2936 (-3.43)	-0.5259 (-2.48)	-0.3139 (-4.63)
CONSTANT	2.2752 (1.25)	2.637 (2.36)	6.035 (2.14)	2.966 (3.30)
Hit Rate	85.71	80.29	83.87	82.6
ρ^2	0.103	0.093	0.164	0.104
$\overline{\rho}^2$	0.088	0.084	0.142	0.1
Sample	245	416	155	816

IVTT : in-vehicle-time
OVTT : out-of-vehicle-time = waiting time + transfer time + walking time

4.3 Findings from SP Analysis

From the viewpoints of the signing condition, the statistical significance and the magnitude of the estimated parameters, some improvement is established in comparison with the RP analysis in spite of the depreciation of ρ^2 and Hit-ratio. Moreover, the range of the magnitude of the parameters including the COST among segment models is not so deviated even under the assumption of "heterogeneity", which may be approaching to the appropriateness of "homogeneity" much closer.

5. ATTEMPT TO COMBINE RP WITH SP DATA

5.1 Description of Combined Estimation

A combined estimation simultaneously using RP and SP data conforms to the following methodology proposed by T, Morikawa et al.

(1) Model Specification

Two different model types are considered: RP and SP models.

RP model:

$$u_{in}^{RP} = \beta' x_{in}^{RP} + \alpha' w_{in}^{RP} + \varepsilon_{in}^{RP} = v_{in}^{RP} + \varepsilon_{in}^{RP}$$

$$i = 1, \ldots, I_n^{RP}, n = 1, \ldots, N^{RP}$$

$$d_n^{RP}(i) = \begin{cases} 1 & \text{if Alternative } i \text{ is choosen by Individual } n \text{ in the RP data} \\ 0 & \text{othewise} \end{cases}$$

SP model:

$$u_{in}^{SP} = \beta' x_{in}^{SP} + \gamma' z_{in}^{SP} + \varepsilon_{in}^{SP} = v_{in}^{SP} + \varepsilon_{in}^{SP}$$

$$i = 1, \ldots, I_n^{SP}, n = 1, \ldots, N^{SP}$$

$$d_{in}^{SP}(i) = \begin{cases} 1 & \text{if Alternative } i \text{ is chosen by Individual } n \text{ in the SP data} \\ 0 & \text{otherwise} \end{cases}$$

where

u_{in} = utility of Alternative i to Individual n,
v_{in} = systematic component of u_{in},
ε_{in} = random component of u_{in},
$d_n(i)$ = choice indicator of Alternative i for Individual n,
x_{in}, w_{in}, z_{in} = choice indicator of Alternative i for Individual n, and
α, β, γ = vectors of unknown parameters.
RP,SP: superscript indicating the data type

The level of random noise in the data sources is represented by the variance of the disturbance term ε. If RP and SP data have different noise levels, this can be expressed by

$$Var\left(\varepsilon_{in}^{RP}\right) = \mu^2 Var\left(\varepsilon_{in}^{SP}\right) \quad \forall i, n \tag{5}$$

If SP data contain more random noise than RP data, μ will lie between 0 and 1. μ is also known to represent the "scale" of the model coefficients.

Assuming independently and identically distributed Gumbel disturbance terms in the RP model, a logit model is obtained with the choice probability given by

$$P_n^{RP}(i) = \frac{\exp\left(v_{in}^{RP}\right)}{\sum_{j=1}^{I_n^{RP}} \exp\left(v_{jn}^{RP}\right)} \tag{6}$$

An independently and identically distributed Gumbel assumption for the SP utility disturbances leads to the following SP logit model, which includes the scale parameter μ:

$$P_n^{SP}(i) = \frac{\exp\left(\mu \cdot v_{in}^{SP}\right)}{\sum_{j=1}^{I_n^{SP}} \exp\left(\mu \cdot v_{jn}^{SP}\right)} \tag{7}$$

(2) Model Estimation

A model is calibrated using the following sequential estimation procedure.

Step 1. Estimate the SP model (Equation 3) using the SP data to obtain $\hat{\mu\beta}$ and $\hat{\mu\gamma}$. Define $y_{in}^{RP} = \mu\beta' x_{in}^{RP}$ and calculate the fitted value $\hat{y}_{in}^{RP} = \hat{\mu\beta}' x_{in}^{RP}$ for the RP observations.

Step 2. Estimate the following RP model with the fitted value $\hat{y}_{in}^{RP}$ included as a variable to obtain $\hat{\lambda}$ and $\hat{\alpha}$:

$$\hat{\mu}_{in}^{RP} = \lambda \hat{y}_{in}^{RP} + \alpha' w_{in}^{RP} + \varepsilon_{in}^{RF} \tag{8}$$

where $\lambda = 1/\mu$. Calculate $\hat{\mu} = 1/\hat{\lambda}$, $\hat{\beta} = \hat{\mu\beta} / \hat{\mu}$, and $\hat{\gamma} = \hat{\mu\gamma} / \hat{\mu}$. The accuracy of $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ can be improved by Step 3.

Step 3. Multiply x^{SP} and z^{SP} by $\hat{\mu}$ to obtain modified SP data set. Pool the RP data and the modified SP data and then estimate the two models jointly to obtain $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$.

5.2 Combined Estimation with RP and SP Data

In order to discuss about the homogeneity regarding the mode choice behavior, a combined model simultaneously using RP and SP data is calibrated. The results of estimation by trip distance are given in Table 8, which shows that μ (scaled parameter) ranges between 0 and 1 successfully to demonstrate the SP bias is larger than the random noise of the RP data. Comparing these results with those from the RP analysis, although ρ^2 of the good-of-fit and Hit-ratio decrease to their value, the tremendous improvement in t-statistics for all the LOS variables especially in the COST is realized. Comparing them with those from the SP analysis, on the other hand, ρ^2 and Hit-ratio increase in their values with slight depreciation in the t-statistics of the COST and the OVTT variables. The modal choice characteristics derived from the combined use of both data of the RP and the SP offers much more statistically accurate information than the single application of RP or SP data, because the combined use may cancel respectively owned disadvantages with each other.

The COST parameter of the medium trip distance model implicitly indicates still slight inconsistency in t-statistics and magnitude with other segment models as shown in Table 8. This is obviously caused by the revealed behavior of the same segment of medium trip distance as shown in Table 5. If the data belonging to the segment of the RP data were refined, much more expected outcome to be able to assume "homogeneity" would be obtained. The overall model is, however, calibrated with the appropriate statistics and the magnitude of estimated parameters in some degree, which suggests that the assumption of "heterogeneity" may not play an important role.

Table 8 Estimated Statistics with RP and SP Data Combined

Variable	Short-dist model	Medium-dist model	Long-dist model	Overall model
COST	-0.48 (-3.77)	-0.0903 (-0.49)	-0.7749 (-3.09)	-0.3106 (-2.65)
IVTT	-0.1013 (-5.01)	-0.0828 (-5.63)	-0.1128 (-3.19)	-0.0828 (-7.36)
OVTT	-0.1106 (-3.41)	-0.1079 (-5.33)	-0.235 (-3.71)	-0.313 (-4.43)
CONSTANT (SP)	0.4874 (0.51)	-0.1192 (-4.80)	0.1367 (1.58)	-0.4719 (-1.59)
CONSTANT (RP)	-0.304 (-3.67)	-0.3623 (-3.40)	-0.1034 (-0.71)	-2.82 (-4.71)
μ (scale parameter)	0.5279 (8.99)	0.3121 (11.46)	0.824 (7.56)	0.3005 (17.04)
Hit Rate	89.1	84.6	87.1	85.9
ρ^2	0.214	0.152	0.231	0.141
$\overline{\rho}^2$	0.204	0.146	0.214	0.138
Sample	378	647	209	1,234

IVTT : in-vehicle-time
OVTT : out-of-vehicle-time = waiting time + transfer time + walking time

6. SUMMARY AND CONCLUSIONS

The followings are the summary of modeling; (1) the RP approach shows that there may exist certain threshold by area and by trip distance category for the modal choice, although estimated parameters of LOS variables, specially COST variable are statistically instable. (2) the SP approach illustrates that the statistical stability and the magnitude of LOS variables appreciates the reasonableness of the mode choice behavior. (3) an analysis using RP and SP data respectively gives some disadvantages in comparison with each other (e.g., low likelihood ratio for the SP approach). (4) the combined approach with RP and SP data improves the aforementioned disadvantages canceled with one another, and provides a better model, which may help the sensitivity analysis to look for what kind of LOS will contribute to the increase in passengers of proposed LRT riders for instance.

This study concludes that the method developed here which requires relatively small sampling size demonstrates one of the effective and practical tools for establishing an appropriate urban transportation planning in developing countries. It must be strongly pointed out to handle not single data set from RP or SP but the combined data set, because the assumption of "homogeneity" may tend to overcome that of "heterogeneity". It will be desired to continue further discussion about this issue in detail aside from the effort to collect much more fruitful data.

ACKNOWLEDGMENT

The study from which this paper is developed was made through a lot of assistance as follows: Professor T. Kurokawa, Tsukuba University, kindly offered financial support from a part of grant provided by Kashima Corporation Research Fund. Mr. H. Felias Jr., University of the Philippines, not only designed the questionnaire form, supervised the survey, discussed and advised about the orientation of the study, but also accomplished a research paper as a first step of the study as listed in "REFERENCES". People in ALMEC in Manila Office conducted and managed the survey. Mr. T. Kawayoke, Mrs. A. Iijima, Mr. K. Okamura, Mr. K. Watanabe, Mr. J. Hisada, Mr. H. Aoki and Mr. K. Mori, who were former students of our Laboratory, helped us in the field of data processing, model building and calculation. The authors would like to note here their contribution respectively and to acknowledge them gratefully.

REFERENCES

Uchiyama, H. and Felias Jr., H. (1995) A study on Metro Manila's modal choice characteristics. 7th WCTR, Sydney, (presenting paper).

Morikawa, T., Ben-Akiva, M. and Yamada, K. (1991) Forecasting intercity rail ridership using revealed preference and stated preference data. Transportation Research Record 1328, 30-35.

Ben-Akiva, M., Morikawa, T. and Shiroishi, F. (1989) Analysis of the reliability of stated preference data in estimating mode choice intentions. 5th WCTR, Vol.4, Yokohama, Japan, 263-277, 1989.

Morikawa, T. (1990) Review and perspective of incorporating stated preference data in travel demand analysis. Journal of Infrastructure Planning and Management, Japan Society of Civil Engineers, No413/IV-12, 9-18, January, 1990.

Morikawa, T and Ben-Akiva, M. (1992) Combined estimation of choice models from RP and SP data. Traffic Engineering, Japan Society of Traffic Engineers, Vol.27, No.4, 21-30, July, 1992.

Kurokawa, T., Ishida, H. and Villaroman, M. B. (1989) The effects of perception and feeling variables on mode-choice behavior in Metro Manila. 5th WCTR, Vol.3, Yokohama, Japan, 701-715, 1989.

Nakamura, R. and Kashima, S. (1989) The basic on mode choice characteristics in Metro Manila. 5th WCTR, Vol.3, Yokohama, Japan, 657-670, 1989.

Tamura, T., Lidasan, H. S. and Kurokawa, T. (1992) Formulation of the modeling framework of panel analysis application in a developing country -Metro Manila: a case study. 6th WCTR, Vol.3, Lyon, France, 475-486, 1992.

Uchiyama, H.and Mohri, Y. (1993) A Study on survey method of urban transportation in developing countries. Proceedings of Infrastructure Planning, Japan Society of Civil Engineers, No16.(1)-2, December, 1993.