A PROCESS MODEL OF TOURISTS' DESTINATION SET CHOICE BASED ON COMBINATION UTILITY OF DESTINATION SET

Kuniaki SASAKI, University of Yamanashi, 4-3-11, Takeda, Kofu, Yamanashi, Japan 4008511, Tel. +81-55-220-8671 E-mail sasaki@yamanashi.ac.jp

Kazuo NISHII University of Marketing and Distribution Science, 3-1 Gakuen-Nishi, Nishi-ku, Kobe, Japan Tel. 078-796-4852, E-mail Kazuo_Nishii@red.umds.ac.jp

ABSTRACT

When planning the transportation in sightseeing area, trip chaining of sightseeing spots is key aspect to understand tourists and attractiveness of sightseeing area. Thinking from the view of travel behaviour analysis, those characteristics are defined as the combination choice of destination and order of visiting. This research focused on those characteristics and proposed a model of process of combination choice of destination. Another aspect we consider in this research is common un-observed factor between sightseeing spots. To consider this problem, we assumed correlations between random terms. This assumption of choice leads to a kind of mixed logit model. As the empirical study of the proposed model, we applied the model to the data of tourism behaviour in Kyoto City in Japan. The result of model estimation showed significant correlation between different three defined categories on sightseeing as the parameter of random term. As a conclusion, we proposed a feasible method to model the combination of destination and to consider the correlation of un-observed factor of sightseeing spots.

Keyword: tourism behaviour, combination choice, heuristic decision rule, mixed logit model

1. INTRODUCTION

The Japanese government started a policy which promoted tourism industry in 2003, named "the Visit Japan Policy". It was coincident with the holiday law change which allowed holidays to fall on Fridays or Mondays, for the purpose of making some weekends longer. In addition, as expressway construction had reached to the rural areas, tourism business is expected to become the main industry in the rural areas in Japan. Since most areas the tourism development plans are aimed at are generally less resourceful, they need to be teamed up with their neighbourhoods by road improvement. However, the Japanese road investment evaluation system has not considered the economic effects of tourists' activities, though one of the important roles of the road improvement is to support tourists' excursion. Therefore, the cost benefit analysis on road investment would become an important factor soon.

Because analyzing tourists' travel behaviour is necessary to evaluate the economic effect, it is necessary to analyze tourists' behaviour.

Tourists usually have excursions, series of trips which links different sightseeing spots. This behaviour specific to tourists often becomes an object of trip analysis of tourists' behaviour, because the transportation policy to solve some problems caused by tourism, such as traffic jams, requires information about OD of tourists within a sightseeing area. To analyze this behaviour, the combinations of destinations tourists travel must be considered. However, forecasting the combination of destinations becomes so difficult because of the large number of possible combination, when the number of destination in an excursion is not limited, but also, we cannot assume that tourists choose their destination combination from all the possible combinations of destinations. One useful method to apply to this case is the sequential choice assumption, which is sometimes used in the activity analysis that considers whole day activities including multi-attributes alternatives. This is a simple and easy model structure, but the information of the already chosen destinations can be used only after the next choice. Therefore, it is hard to consider the whole utility of excursion. Another possible model is a hierarchical choice structure, but this is not practical because it has to take so many possible set into account if the number of destinations in a trip is not few.

Besides, the utility of multi-destination tourists' excursion cannot be a sum of the utilities of each destination because the tourists' tastes generally are constant throughout an excursion: Adding one bustling leisure land to a tour visiting heritage sites may not increase the total utility of the tour, because tourists who prefer quiet heritage sites usually do not like bustling places. If there is a feasible method that can calculate the utility of multi-destination tour, analyzing excursion behaviour would be more effective for policy analysis.

2. REVIEW OF RELATED RESEARCHES

In this study, we study tourists' behaviour focused on the combination of destination and the aggregate utility of the combination. The destination choice of tourist is intrinsically difficult to analyze because of the following four aspects (Okamoto et al, 1995)

1) Rare behaviour: the number of the tour per year is extremely less. For example, in Japan, the average number of sightseeing trip in a year per person is 1.05 in 2008.
2) Non regularity: Touring is not a regular behaviour, but is affected by some factors such as seasonal factors, experiences and information. Therefore the choice analysis of this behaviour has to consider more factors than regular behaviour.
3) Heterogeneity: the motivations and the styles of tours are diverse. The tourists who visit a certain spot choose it based on different motivations and different styles. That is, decision making in touring is heterogeneous.
4) Geographic Scale: Geographic scale of sightseeing is different from the ordinary trips. Besides, the scale of trips in a sightseeing tour is not the same between the access trips and the excursing trips.
Researches on sightseeing behaviour in Japan have taken these aspects into account.

The analytical framework of tourist behaviour was proposed by Moutinho (1987). It describes both pre-purchase and post-purchase behaviour of tourists. Regarding the behaviour aspect on destination choice in particular, Parsons and Needelman (1992) and Parsons and Kealy (1992) proposed methods to specify choice set by aggregation of some alternatives. Um and Crompton (1991) proposed two set, the evoked sets and the awareness sets with internal and external factors, on the destination choice on leisure travel. Crompton (1992) had developed the taxonomy of vacation destination choice sets, and published a series of researches about destination choice since then. A set of propositions relating to choice sets was offered by Crompton and Ankomah (1993). Some of them are listed below.

Proposition 1: The respondents' vacation destination priorities correspond positively with the order in which they are mentioned in unaided recall questions.
Proposition 2: The average number of destinations that individuals will seriously consider in their late consideration sets in making vacation decisions will not exceed four.
Proposition 3: the perceived importance and/or the perceived risk of the destination decision will not have effect on the size of the late consideration set.
Proposition 4: A destination which a potential tourist makes more effort in getting information is more likely to be in the late consideration sets unselected at the final choice.

Crompton et al (1999) reported the tests result of the four propositions or corollaries. The basic concept of their analysis is focused on the latent set of destination, such as the evoked set, the consideration set and the choice set. Decrop (1999) extended the latent set, which are the available set, the awareness set, the dream set, the evoked set, the exclusion set, the surrogate set, the unavailable set. These researches mainly treated the choice of destination itself rather than choice set. The study of destination combination in this paper is similar to the choice-set generation, since the model of choice set generates a set of alternatives from the universal choice set. Therefore we review the choice set models next.

The modelling for the choice set generation in the discrete choice scheme was commenced by Manski (1977). Manski suggested a basic concept of the choice behaviour analysis based on an obscure choice-set model. That is PCS (Probabilistic Choice Set) model which became common in analyzing the choice set. Many theories and models have taken the choice-sets into account implicitly or explicitly to improve travel behaviour forecasting. Many of the proposed models assume the probabilistic choice set because the indicators and information about choice set are hard to acquire. One trial by simple assumption is Dogit model by Gaudry and Daganais (1979), which assumes two classes. An improved Dogit model is proposed by Swait and Ben-Akiva (1987). They named it PLC (Parameterized Logit Capacity) model. Ben-Akiva and Boccara (1995) assumed a latent choice-set with some indicators of alternative availability. Morikawa (1995) proposed choice set modelling by using a paired comparison of each alternative of destination choice. Haab and Hicks (1997) treated choice set as an evoked set of alternatives whose probability follows normal distribution. Ben-Akiva et al (1997) suggested a general modelling scheme of choice set model by latent class approach. As Thill (1992) pointed out, probabilistic model would be better applicable method for choice set modelling, but, these probabilistic models are hard to apply to our case because our final goal is to specify the combination of destination.

Some researches on choice set focus on information processing ability of human being. Meyer (1980) incorporated information gathering process in the context of destination. In Meyer's model, when the adding cost of an alternative into choice set is higher than the expected utility of the alternative, that alternative will not be included to choice set. This is a kind of learning process in destination choice. Fotheringham (1988) suggested a competing destination model for store choice based on information-processing strategy. His assumption is two stage choices, one chooses a cluster of stores, then choose a store from that cluster. Other than that, theoretical models such as EBA (Tverskiy, 1972) or some non-compensating processes are proposed, but there is no model practical enough to be applied to analyzing choice set. Crotts (1999) suggested the application of pre-purchase information search process to tourism, which investigates the information acquisition process of a decision maker. This is one of the process models for tourists' decision making.

When thinking about combination utility, it will be a problem of similarity of alternatives because some components of alternative are the same. That is, IIA problem in logit model scheme cannot be omitted in combination utility. Thus, We review simply the manipulation of similarity of alternative.

After proposing of GEV model (McFadden, 1978), some models were developed to relax the constraint of MNL, such as Nested Logit model (Ben-Akiva and Lerman, 1985), Cross Nested Logit model (Small, 1987), Paried-Combinational model (Chu, 1989), Heteroscedastic Extreme Value model (Bhat, 1995), and Generalized Logit Model (Chieh-Hua and Koppelman, 2001). Mixed Logit model (McFadden and Train, 1997) became popular method to consider the correlation of random errors, because it is possible to provide a priori structure of correlation structure in the model as Train (1998) applied this model for destination choice of fishing-site. The similarity of combination utility in this study is easy to specify the correlation structure because the component of alternative is explicit. Therefore, we will use mixed logit model in this study.

3. MODELLING FRAMEWORK

1) Choice Model of Combination of destination

As mentioned in the chapter 1, to analyze excursion behaviour it is necessary to model the destination combination. Because the number of the combination of destination will be big when the number of destination is not small, ordinary approach such as multinomial choice model is hard to apply. For example, when the destinations in universal choice set are 10, the number of possible combination without any constraint on the matching becomes 1023. The chance that all the decision makers choose the same combination as the most desirable is little. Then we assume that each decision maker takes a process which will lead to a satisfactory goal in choosing destination set as Sasaki et al (2004) showed. That is a repetitive process to choose the most desirable combination among resembling combinations started from a null set. In other words, decision makers choose the most preferable destination from all the destinations first, and then add another one in order to explore more desirable choice set. They repeat this process until the chosen combination is the same one as the precedent choice. These choices are based on an assumed utility function at each process. Firuge-1 shows this process by a flowchart. This process is feasible because this surely yields one destination combination and the number of alternative in each process never exceeds the number of all the destinations. However, the whole process is hard to observe, and what we can observe is only the final combination of destination and all the destinations, but not choice on each process. Therefore, we assume next: "the utility function is common to all processes". This assumption enables us to estimate the parameter of utility function only by final choice.

Here, we define this process by mathematical notation. We define the utility function of combination of the destination as equation 1) when we notate the combination of destination as c. Note that we omitted suffixes for individuals and alternatives for simplicity.

$$U_c = \boldsymbol{\beta}\mathbf{x}_c + \varepsilon_c \qquad 1)$$

Where
U_c: the utility of combination of destination
$\boldsymbol{\beta}$: a vector of unknown parameters in the utility function
$\mathbf{x}_c$: a vector of attributes of combination for destination combination *c*
ε_c: error component of destination combination *c*

When we assume the error component as I.I.D. Gumbell, the choice probability of combination *C* is derived as equation 2)

$$P(C) = \frac{\exp(\mu\, V_C)}{\sum_{j \in J} \exp(\mu\, V_{C+j})} \qquad 2)$$

Where
C: the final chosen combination of destination
J: The set of destinations which are not included to combination *C*
μ: scale parameter of Gumbell distribution
C+j: A combination of destinations which added one more destination from *J* over the final chosen combination *C*

Suppose there are four destinations, {A, B, C, D}, and suppose the actual choice of destination combination was {B, C}, the initial process becomes choice from four alternatives {A} {B} {C} {D}, and the final process is assumed as a choice from three alternatives {B,C} {A, B, C} and {B, C, D}. Parameter of utility function is estimated only from the final process. This process model offers feasible solution to this combination choice because the number of the alternative combination is always less than the number of destinations. With the ordinary approach without any constraint, the choices are fifteen, such as {A}, {B}, {C}. {D}, {A, B}, {A, C} ... {A, B, C, D}. When the destinations are around five, this approach is still feasible but because the alternatives increase like power series, the efficiency of parameter estimation decreases in contrast.

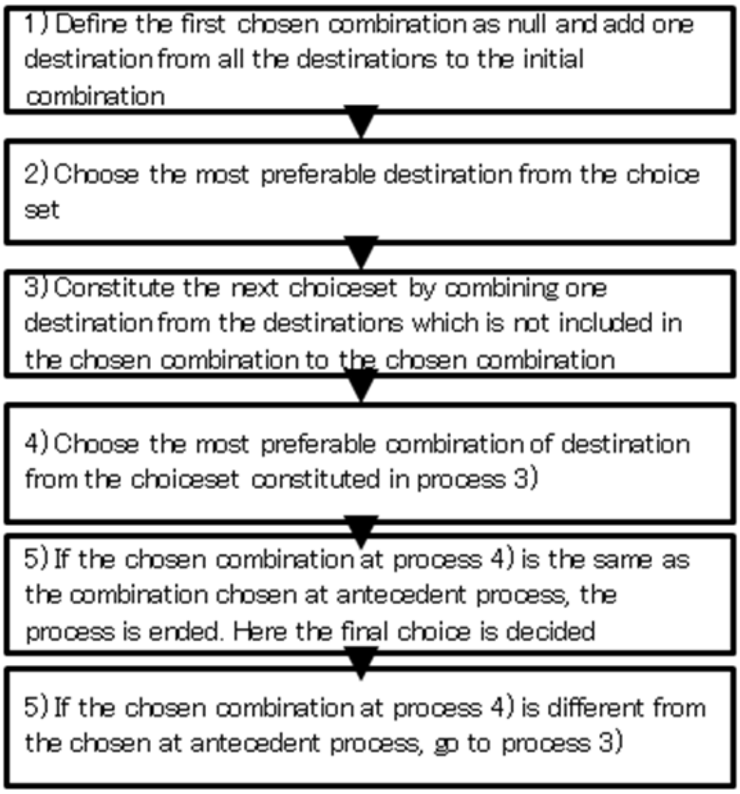

Figure-1 The flowchart of the assumed process

(2) Combination utility

The model we proposed does not consider the feature of combination utility. As mentioned in the first chapter, the combination of destination would yield synergistic effects or anti-synergistic effects, especially on sightseeing excursion because the taste heterogeneity strongly affects on this behaviour. Then, we reconsider the utility function by incorporating the effect of the destination combination. We modify the equation 1) to consider those effects. Suppose the utility function of combination of destination *i* and *j* as equation 3).

$$U_{ij} = \boldsymbol{\beta}\mathbf{x}_{ij} + \varepsilon_{ij} \qquad 3)$$

Here, the attribute vector can be decomposed as equation 4)

$$\mathbf{x}_{ij} = \mathbf{x}_i + \mathbf{x}_j + \mathbf{x}_{ij} + \alpha_{ij} \qquad 4)$$

Where

$\mathbf{x}_i, \mathbf{x}_j$: vector of attributes of destination *i* and *j*, respectively

$\mathbf{x}_{ij}$: vector of attributes associated with destination i and j, such as distance between two destinations

α_{ij} : interaction component associated with the combination of destination i and j

Here one of the unknown parameter in $\boldsymbol{\beta}$ as γ , the utility function is rewritten as equation 5)

$$U_{ij} = \boldsymbol{\beta}(\mathbf{x}_i + \mathbf{x}_j + \mathbf{x}_{ij}) + \gamma\alpha_{ij} + \varepsilon_{ij} \qquad 5)$$

Where

γ_{ij} : unknown parameter for interaction component

Because observing component interaction is generally difficult, α_{ij} is treated as one unit parameter at estimation. If there are a certain number of destinations, the number of unknown parameter for interaction term increases so much. Then, to reduce the parameters to estimate, we assume: "the destinations can be categorized and the interactions between destinations from the same categories are common ". That is, for two destinations i and j which belong to the category m and n respectively, interaction α_{mn} does work to that case. From this assumption, equation 5) is modified as below.

$$U_{ij} = \boldsymbol{\beta}(\mathbf{x}_i + \mathbf{x}_j + \mathbf{x}_{ij}) + \mathbf{d}_{ij}\boldsymbol{\alpha}_{m,n} + \varepsilon_{ij} \qquad 6)$$

Where

$\mathbf{d}_{ij}$: a row vector of dummy variables, 1: if destination *i* and *j* belong to category *m* and *n* respectively, 0; otherwise

$\boldsymbol{\alpha}_{m,n}$: a column vector of interaction components, in which each variable represents the interaction of category *m* and *n*

This is regarded as a simple logit model with a linear utility function with category dummy variable if the interaction component is deterministic for each category. Meanwhile, the interaction components among destinations were summarized as interaction among destination categories along the assumption for simplicity, the deterministic characteristic for this component seems not to be adequate. Relaxing this constraint, it is desirable to allow a distribution of interaction component in order to take account of the diverse destinations in the same category. Thus, the model structure is not multinomial logit but becomes mixed-logit structure (McFadden and Train, 1997) as mentioned in Chapter 2

4. EMPIRICAL ANALYSIS

1) data used

The data we used in this empirical study is a survey on excursion of tourists at KeiHanShin (Kyoto-Osaka-Kobe) Metropolitan area in Japan in 2005. This survey is a part of comprehensive survey whose purpose is to grasp the traffic flow on weekend in KeiHanShin metropolitan area. It was conducted on a weekend in October 2005. The form of survey was onsite survey at twenty seven famous sightseeing spots in KeiHanShin area by hand-on and the mail-return. The main items in the questionnaire were categorized into three parts. The first one is the style of tour and itinerary, such as "the number of the participant in the same tour" and "a day trip or a multi-days trip". The second one is the attributes of the excursion, such as "the visited tourism facilities", "the purpose of the visit on each facility" and "transport mode to the facility". The last part is individual attributes such as gender and age.

We used a part of this data which was distributed at tourism facility in Kyoto city area, because the KeiHanShin area is too large and the whole data involve too many samples to analyze. Kyoto is a Japan's old capital and is the most attractive city for the tourists in and to Japan, and has a lot of old temples and old shrines. The facilities in which the survey was distributed in Kyoto area were six spots such as "Byo-Do-In", "Ko-Myo-Ji", "Shijo-Karasuma" and so on. The detail of this survey is shown in the report of the KeiHanShin Association of Transportation Planning published in 2007.

Since the data has 1729 samples from wide-ranging tourists, the visited sight spots are numerous. Then we summarized these facilities into twenty five typical sightseeing zones by using some published guidebooks and some sightseeing information web-site. These zones are shown in Table-1. Because this survey was on-site, the sampling does not represent the population of tourist behaviour. We assign the expanding rate to each sample by sampled probability which was calculated by complementary survey of counting tourist. The number of expanded sample is 57505, which means that the average sampling rate is about 3%. We categorized these twenty five zones into three categories shown below. The number of possible combination is

Table-1 the zones categorized by its nature

Downtown	Nakagyo, Kyoto Station
Shrine and Temple	Kinkaku temple , Kami-Gamo Shrine, Ginkaku-temple-Heian-Shrine, Shisen-do temple, Kitano Tenmangu Shrine, Old palace, Toh Temple, Nijo castle, Joyo-city, Koryuji
Natural beauty	Nagaoka old capital, Uji river, Ohara-Kurama, Kiyomizu temple, Yasaka shrine, Higashi Hongan temple, Yamashina, Fushimi, Zenpo temple, Arashi-yama, Kameoka, Yawata-Kumiyama, Katsura

1) downtown: zones at the centre of the city
2) shrine and temple: zones located in the city area with old temples or old shrines
3) natural beauty: zones located in the suburban area with beautiful gardens or natural tourism resources. Some of those gardens and nature are in the famous temple or shrine

In the last category, some shrine and temples are included, because those have beautiful gardens and are on hillside. The classification of these two zones relies on the descriptions in the guidebooks of those temples and shrines. If the main topic of a temple or shrine is on their gardens or hills, we classified it into the third category.

2) Basic statistics of the data

In this section, we show the basic statistical description of the data we used. Because the data was the choice-based survey data, we show the statistics of the expanded data which is adjusted to the population of the tourists at that day.

First, we show the top ten most visited facilities in each sightseeing zone and their share in the expanded population in Table-2. This shows that most samples here visited the surroundings of the Kyoto Station and Nakagyo-ku, which characterizes this data. The reason why the zone 2 was the most visited one might be visitors' travel mode. More than 80% of samples used public transportation, particularly trains as access mode to Kyoto area. These two zones have a train terminal inside respectively, in addition to the attractive facilities to visitors. Tourists visited these zones before and after taking train for access and egress to Kyoto area. Moreover, this survey focuses on the visitors who visited attractive spots at six zones; this means that the main purpose of a part of sample may be shopping at downtown area. The Uji-river and Nagaoka, an old capital, where we distributed questionnaire, attracted a large number of visitors.

Table-2 The distribution of samples' visiting zones

Rank	Sightseeing zone	The number of visitors (Expanded)	Ratio per total	Ratio per sample
1	Kyoto Station	22507	23.5%	39.1%
2	Nakagyo-ku	21387	22.3%	37.2%
3	Uji-river	8585	9.0%	14.9%
4	Kiyomizu temple	6995	7.3%	12.2%
5	Nagaoka old capital	6670	7.0%	11.6%
6	Ginkaku and Heian	6389	6.7%	11.1%
7	Nijo castle	3705	3.9%	6.4%
8	Arashiyama	3570	3.7%	6.2%
9	Yasaka shrine	2287	2.4%	4.0%
10	Higashi Hongan temple	2015	2.1%	3.5%
	Total number of visited zone	95805	Number of sample	57505

And then we show the distribution of the number of visited zones by each sample in Figure 2, where the horizontal axis is the number of visited zones and the vertical axis is the number of expanded sample. From this figure, most of the samples visited one zone only, which takes about fifty six percent of the totals. Since there are 325 possible combinations of destinations, even though the number of visit destination is limited to two, it is difficult to estimate a significant model of combination choice from universal set. Table-3 shows the number of the samples who toured more than two zones, which include different category. The combination of "shrine and temple" and "natural beauty" is lower than other combination. Table-4 shows the number of samples who toured more than two zones in the same category. This shows that the number of samples who toured "shrines and temples" multiple times is lower than other combinations, while the repeated combination of "downtown" and "natural beauty" is more popular.

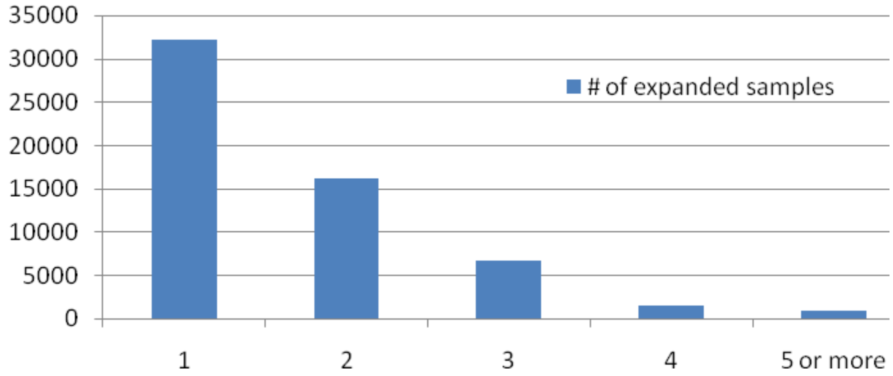

Figure-2 the number of samples over the number of visited zones

Tabel-3 the number of samples who toured Combination of different zones

Shrine and Temple + Downtown	Shrine and Temple + Natural beauty	Natural beauty + Downtown	Total
8472(36%)	6304(27%)	8631(37%)	23407

Tabel-4 the number of samples who toured Combination of same zones

Shrine and Temple	Natural Beauty	Downtown	Total
1895(12%)	6172(40%)	7524(48%)	15591

3) Model estimation

On the proposed model which assumes an exploring process over combination of destination, we assume zero for the combination of the same zone to standardize the interaction. Consequently, we estimate three $\alpha_{m,n}$ in equation 6) as difference from the combination of the same zones. There are two possible approaches, one is deterministic $\alpha_{m,n}$ and the other is probabilistic $\alpha_{m,n}$ approach, as mentioned in the previous chapter. Deterministic approach is quite simple logit model, but the probabilistic approach needs some additional assumptions to estimate. We assume that the distribution of $\alpha_{m,n}$ is normal distribution and the expectation value of $\alpha_{m,n}$ must be a parameter. The larger is the variance of the interaction, the larger the expectation of interaction becomes, because the source of variance is from the heterogeneity of summarization of three categories. We formulate this assumption as shown in equation 7).

$$\alpha_{m,n} = \lambda_{mn}\zeta \qquad 7)$$

Where

λ_{mn} : unknown parameter which express the expectation of interaction component

ζ : a random variable ~N(1,1)

The estimated results of the utility parameters are shown in Table-5. Right side of the table is estimation result of deterministic interaction model and the left side is the probabilistic interaction model. The explanatory variables are explained in Table-6. ST*DT is the parameter for the interaction component between "Shrines and Temples" and "Downtown", ST*NB is the parameter for the interaction component between "Shrine and Temple" and "Natural Beauty", and NB*DT is the parameter for the interaction component between "Natural Beauty" and "Downtown". These three variables correspond to the dummy variable of interaction components.

Table-5 Estimation results of the utility parameters

Variables	Probabilistic interaction model	Deterministic interaction model
ANSF	0.0830 (15.9)	0.0845 (16.0)
ANP	-0.0202 (-6.7)	-0.0207 (-6.8)
ANSS	-1.17 (-18.3)	-1.05 (-15.7)
Distance	-0.123 (-21.2)	-0.123 (-21.1)
ST*DT	-0.152 (-1.3)	-0.250 (-2.1)
ST*NB	-0.345 (-3.2)	-0.477 (-4.4)
NB*DT	-0.696 (-5.9)	-0.868 (-8.3)
Adjusted ρ2	0.462	0.464
number of Expanded samples	57505	57505

( ) t-statistics

Table-6 Definition of explanatory variables

ANSF	The average number of tourism facilities including sightseeing spots in some guidebooks in a combination of zones
ANP	Average number of sum of the pages of each zone in a combination of zone, in some guidebooks
ANSS	Average number of sightseeing spots on each zone of the combination of zones defined in some guidebooks
Distance	Total distance from the average of each centroid of zone to each centroid of zone
ST*DT	1: if “Shrine and Temple” and “Downtown” zone is included in a combination of zone, 0: otherwise
ST*NB	1: if “Shrine and Temple” and “Natural beauty” zone is included in a combination of zone, 0: otherwise
NB*DT	1: if “Natural beauty” and “Downtown” zone is included in a combination of zone, 0: otherwise

Because both the indicators of goodness-of-fit of the models are not so different and the estimates of attributes of combination of zones are also not so different, we comment on the result simultaneously. The significant positive estimate of ANSF parameter means that a combination of zone is more preferred, when the average number of tourism facilities increases. That is, if the number of tourism facilities in the added zone is larger than the average number of tourism facilities in the original combination of zones, the attractiveness of the new combination will increase, but will decrease when the number of tourism facilities in the added zone is less than the average number of that of the original combination. On the other hand, the estimates of ANP and ANSS are significant negative. If the number of pages in guidebooks and that of sightseeing spots in the added zone is larger than the average of that of the original combination of zones, the attractiveness of the new combination increases, but decreases when the number of pages in guidebooks and that of sightseeing spots in the added zone is less than the average number of that of the original combination. This is a little bit strange but reasonable, because most of the samples visited "Downtown" zone where the pages in guidebooks and the sightseeing spots are not many but many facilities for tourist such as shopping complex including souvenir shops are located. In general, when adding one zone to "Downtown", ANP and ANSS will increase but ANSF will decrease. The results that the most visited zone is just one and that "Downtown" category is preferred most imply that these estimated results are reasonable. This implication is derived from such data characteristics. The result that parameter for "Distance" was estimated as significant negative means that the combination of zones whose locations are not close is not preferred. This is intuitively reasonable.

About interaction components, the difference of estimated parameter is not significant and the value order of each parameter remains the same. However, the absolute value of each interaction parameter in probabilistic model is relatively smaller than that of deterministic model. Besides that the t-statistics are relatively small and ST*DT became insignificant. Basically, the degree of freedom is higher in probabilistic model. Consequently, the deterministic model possibly estimates the effect of interaction excessively. One of conceivable reason is asymmetry of distribution of interaction component, because if the distribution is symmetric, the expectation would be the same. Besides, the assumption of the proportional variance as the expectation would be another reason for becoming insignificant.
Focusing on the estimated parameter for interaction, all the parameters are negative. That is, the combination of zones which belong to different category is not preferred to that of the same category. Especially, the interaction component of NB*DT has the biggest absolute value of the parameter. The second biggest absolute value is ST*NB and the smallest is ST*DT. From this result, some implications are derived. For example, when suggesting a route for sightseer, those routes should consist of similar zones. Furthermore, considering the revitalization of city centre in local city in Japan by tourism the model route with "Downtown" should be combined with "Shrines and Temples" type zones than "Natural beauty" zone. When sightseer is not familiar with the detail of visit region, they will rely on the information of route and destination in guidebooks and website. If showing model route in them, interaction should be considered.

5. CONCLUSION AND DISCUSSION

In this study, we focused on destination combination of excursions to analyze travel behaviour of tourists. In particular, we adopted a kind of heuristic model which formulates a process of exploring a satisfying combination from a number of destinations. Besides, we introduced interaction effects to that model and compared probabilistic component and deterministic component. The result of the estimation of these models was similar, that is, the interactions between different zones in comparison to the interactions between the same zones do not increase utility. One possible bias to the deterministic model is over-estimate of the interaction effect, when the two models are compared.

True, our methodology has some essential problems. For example, because the distribution of interaction relies on the heterogeneity of destinations in the same category, it is better to re-examine the method to categorize the destinations by considering the heterogeneity. Heterogeneity of zone must be substantial, however, on the phase of decision making, which is affected by tourists' individual image of zone. It is necessary to study about image of zones the tourists have. Since incorporating transportation related variables into this empirical analysis would enhance the value of this model tremendously, we will take such variables as the next element to deal with. Through application to other data, we are going to brush up the methodology to be applicable to practical use. Even though there are still problems, we conclude that our proposed methodology can be applicable to analyze tourists' excursion behaviour on a number of destinations by considering interaction of destinations.

REFERENCES

Ben-Akiva, M. and B. Boccara (1995). Discrete choice models with latent choice sets, International Journal of Research in Marketing, Vol.12, PP.9-24.

Ben-Akiva M. and S. R. Lerman, (1985). Discrete choice analysis: theory and allocation to travel demand, MIT Press, Cambridge, MA.

Ben-Akiva M., D. McFadden, M. Abe, U. Bockenholt, D. Bolduc, D. Gopinath, T. Morikawa, V. Ramaswamy, V. Rao, D. Revelt and D. Steinberg (1997). Modeling methods for discrete choice analysis, Marketing Letters, Vol.8, pp.273-286.

Bhat, C. R. (1995). A heteroscedastic extreme value model of intercity travel mode choice, Transportation Research Part B, Vol.29, pp.471-483.

Wen, C. and F. S. Koppelman (2001). The generalized nested logit model, Transportation Research Part B, 35, pp.627-641.

Chu, C. (1989). A paired combinatorial logit model for travel demand analysis, Proceedings of the Fifth World Conference on Transportation Research, Vol.4, pp.295-309.

Crompton (1992). Structure of vacation destination choice sets, Annals of Tourism Research, 19(3), pp.420-434.

Crompton and Ankomah (1993). Choice set propositions in destination decisions, Annals of Tourism Research, Vol. 20, Pages 461-476

Crompton J. L, C. Botha and S. Kim (1999). Testing selected choice propositions, Annals of Tourism Research, Vol. 26, pp. 210-213.

Crotts, J. C. (1999). Consumer decision making and pre-purchase information search, Consumer Behaviour in Travel and Tourism, Pizam and Mansfeld eds., Haworth Press, NY, pp.149-168.

Decrop, A. (1999). Tourists’ decision-making and behaviour processes, Consumer Behaviour in Travel and Tourism, Pizam and Mansfeld eds., Haworth Press, NY, pp.103-133.

Fotheringham, A. S. (1988). Consumer store choice and choice set definition, Marketing Science, Vol. 7, pp. 299-310.

Gaudry, M. and M. Daganais (1979). The dogit model, Transportation Research Part B, Vol.13, pp.105-111.

Haab T. C. and R. L. Hicks (1997). Accounting for choice set endogeneity in random utility model of recreation demand, Journal of Environmental Economics and Management, Vol. 34, pp. 127-147.

Manski, C. F. (1977). The structure of random utility models, Theory and Decision, Vol. 8, pp.229-254.

McFadden, D. (1978). Modeling the choice of residential location, Transportation Research Record, 672, pp.72-77.

McFadden, D. and K. Train (2000). Mixed MNL models for discrete response, Journal of Applied Econometrics, Vol. 15, pp. 447-470.

Meyer, R. (1980). Theory of destination choice set formation under informational constraints, Transportation Research Record, Vol. 750, pp.6-12.

Morikawa, T. (1995). A hybrid probabilistic choice set model with compensatory and noncompensatory choice rules, Proceedings of 7th World Conference on Transportation Research, Vol. 1, pp.317-325.

Moutinho, L. (1987). Consumer behaviour in tourism, European Journal of Marketing, Vol. 21, pp.2-44.

Okamoto, N., T. Yai, S. Morichi and T. Nishimura (1995). A study on regional difference of recreational destination choice behaviour, Journal of the Eastern Asia Society for Transportation Studies, Vol.1, pp.351-360.

Parsons, G. R. and M. Kealy (1992). Randomly drawn opportunity sets in a random utility model of lake recreation, Land Economics, Vol. 68, pp.93-106.

Parsons G. R. and M. S. Needelman (1992). Site aggregation in a random utility model of recreation, Land Economics, Vol. 68, pp.418-433.

Sasaki, K., K. Nishii and I. Nishino (2004). A Study on Micro-Simulation of Tour Behaviour and Evaluation of the Simulation System based on a Classification of Travel Patterns, Presented at World Conference on Transportation Research 2004, Istanbul.

Small, K. (1987). A discrete choice model for ordered alternatives, Econometrica, Vol.55, pp.409-424.

Swait, J. and M. Ben-Akiva (1987). Incorporating random constraints in discrete models of choice set generation, Transportation Research Part B, Vol.21, pp.91-102.

Thill, J. C. (1992). Choice set formation for destination choice modelling, Progress in Human Geography, Col. 16, pp.361-382.

Train, K. (1998). Recreation demand models with taste differences over people, Land Economics, Vol. 74, pp.230-239.

Tversky, A. (1972). Elimination by aspects: A theory of choice, Psychological Review, Vol. 79, pp.281-299.

Um, S. and J. L. Crompton (1991). Development of pleasure travel attitude dimensions, Annals of Tourism Research, Vol.18, 374-378.