

## RECREATIONAL DEMAND MODEL AS A MIXTURE OF INTERRELATED TRAVEL BEHAVIORS

Daisuke FUKUDA  
Graduate Student  
Department of Civil Engineering  
University of Tokyo  
7-3-1, Hongo, Bunkyo-ku, Tokyo,  
113-8656 Japan  
Tel: +81-3-5841-6129  
E-mail: fukuda@planner.t.u-tokyo.ac.jp

Shigeru MORICHI  
Professor  
Department of Civil Engineering  
University of Tokyo  
7-3-1, Hongo, Bunkyo-ku, Tokyo,  
113-8656 Japan  
Tel: +81-3-5841-6125  
E-mail: smorichi@planner.t.u-tokyo.ac.jp

**Abstract:** This paper develops a framework of recreational travel behavior model that considers the interrelations among different choices, such as destination choice and travel mode choice. To specify the preference interrelations, the concept of "product bundling" will be introduced to the discrete choice model. This framework can be useful in formulating marketing strategies for recreational sites, because it can make the suggestions about what kind of combinations of recreational components are more preferred by consumers. Empirical analyses were employed to the behavioral survey data extracted from the recreational travel survey in Nara-Prefecture, Japan.

**Key Words:** Discrete Choice, Recreational Demand Model, Product Bundling, Market Basket Analysis, Multiple-Category Decision-Making

### 1. INTRODUCTION

Demand model for recreational travel is usually composed of several choice models representing different travel choice elements, such as departure time, destination, travel mode, route, and excursion. It assumes that travelers select a combination of these components. In traditional travel demand models, the choice decision for each component is assumed to be independent of others, and researchers paid very little attention to the possible interrelationships among decision components.

However, in real world there can exist specific or implicit interrelations among these choice variables. For example, travel agencies often offer various types of recreational package tours composed of a particular combination of travel decision components. Usually, the price of travel package tours is less than the simple addition of the price of individual travel components included in the package. It implies that there exists an interdependence among price of decision components as reflected in the discount of prices. Another example which is not so explicit but very intuitive is about the interrelationship between the attractiveness of a particular destination and travel mode. Travelers tend to think that they will go to particular recreational destinations by particular travel modes, because they may feel the premium of the attractiveness of recreation created by the particular combination of destination and travel mode.

With this background, this paper develops a framework of recreational demand model that considers specifically the interrelations among different choices, such as destination choice and travel mode choice. Specifying the interdependence between the choices of different categories is probably not an easy task, even if we use a powerful and sophisticated modeling framework such as nested logit. It is because there are no interrelations between different levels of nests in nested logit model. In order to specify the interrelations, the concept of "product bundling" or "market basket analysis" will be introduced to recreational demand model within the context of discrete choice modeling. This framework could be useful in formulating the marketing strategies for recreational sites, because it can make the suggestions about what kind of combinations of recreational components are more attractive to consumers. Finally, the framework is empirically tested by using behavioral data extracted from the recreational travel survey in Nara Prefecture, Japan.

We make a brief review on the previous recreational travel behavior analysis and the marketing literature related to the interdependence among choice behaviors in section 2. In

section 3, the modeling framework of interrelated discrete choice is presented. Finally, an empirical analysis is carried out with the recreational travel survey data in order to reconfirm the interdependencies among different choices in section 4.

## 2. RESEARCH REVIEW AND POSITON OF OUR STUDY

In the context of travel behavior analysis, nested logit (NL) model is one of the most widely used models, which consider the multidimensional choice behavior (e.g. Morikawa *et al.*, 1995). Usually, in the NL framework, alternatives belonging to each nest stage are composed of a single choice category only, that is, the first stage for travel mode choice, the second for trip frequency, the third for trip destination and so on. This is mainly because of the convenience for the identification of tree structure. The NL structure, however, has several limitations. First, the correlation structure, which specifies the interdependence between choice alternatives, exists not among the alternatives belonging to different nest stages, but among the alternatives belonging to the same nest stage. Hence, it admits the interdependence only within the same choice categories such as travel mode, and does not allow for the interdependence between travel mode and destination. Second, there may be potentially many different choice structures and the researcher may have trouble in deciding which nest structure is best empirically. Third, even if the nest structure is specified such as the upper level for destinations and the lower level for travel modes, it only describes that the travel modes for the same destinations are more similar than the travel modes for the different destinations. The NL framework, hence, cannot specify the interdependence between different choice categories in general. In the same way, even if the extended models from generalized extreme value (GEV) family, such as paired combinatorial logit, cross-nested logit and generalized nested logit (e.g. Koppelman and Sethi, 2000) are employed, the structures of these models cannot specify the interdependence between different categories. They only allow for the interdependence between 'elemental alternatives', which are composed of the combinations of destination alternatives and travel mode alternatives. Moreover, the cross effects beyond different categories cannot be incorporated in such a framework.

Some models succeeded in avoiding such multidimensional choice problems. Tay *et al.* (1996) developed a portfolio model of recreational trips, which included destination, duration and frequency, assuming that travelers choose one out of a set of trip portfolios which are composed of several choice categories. In this framework, Train *et al.* (1987) proposed a portfolio model of local telephone services. These models, however, do not specify the interdependence between different choice categories.

On the other hand, marketing science literature have many research perspectives on so called "multiple-category decision-making", in which the choice of one product is affected by the presence of other products belonging to different categories (e.g. Russell *et al.*, 1997; Russell *et al.*, 1999). According to Russell *et al.* (1999), there are three types of famous cross-category choice dependence: cross-category consideration, cross-category learning, and product bundling. Especially, product bundling is said to be the most important concept, which has also been employed in our study. Product bundling is defined as a choice process, which results in the selection of two or more non-substitutable products (Russell *et al.*, 1999). This definition implies that in the multiple category choice situations, such as books and movies as alternative entertainment choices, choices in one category alter the utility of choices in other categories. Moreover, it is said that bundling strategy of existing products can result in cost savings due to the presence of economics of scale. Hence, there exists a considerable literature on optimal bundle pricing policies from the perspective of the supply side (e.g. Hanson and Martin, 1990; Chuang and Sirbu, 1999).

For the demand side, bundling can be used as an effective tool for extracting consumer surplus. If we get to know the more preferable combinations of choice categories and non-preferable ones, we can design an ideal fixed bundling that produces more consumer surplus. For example, it is probably not a good idea to offer simultaneous discounts on the more preferable combinations of products, if they tend to be bought together. Instead, discounted one would pull in sales of the other. The understanding of the preference interdependence between different categories is indispensable for such goal.

The concept "Market Basket Analysis" may be useful for inspecting the customers' preference

structures on several products. The analysis mainly focuses on predicting the choice of a bundle (or basket) of multiple-category products on the shopping trip. The main feature of market basket analysis is to incorporate the interdependence in demand relationships across the categories in the final basket when cashing. Recently, several researches incorporating the market basket analysis into the discrete choice framework can be seen. Ben-Akiva and Gershensfeld (1998) presented the modeling framework based on nested logit model and applied it to enhanced custom-calling features for residential telephone service choice. Manchanda *et al.* (1999) proposed a multivariate probit model allowing for simultaneous, interdependent choice of many items and formulated the utility functions incorporating the affection of the pricing and promotional change in one category to other categories. They concluded that not accounting for such factors simultaneously could lead to erroneous inferences. More recently, Russell and Peterson (2000) proposed the multivariate logistic market basket model assuming that the researcher can specify the probability that a consumer chooses one category in the basket, given an information on the actual choice outcomes in all other choice categories. They emphasized the ease of computation of the proposed model.

Among these three market basket models, the model in Ben-Akiva and Gershensfeld (1998) cannot identify the interrelated preference structure endogenously, since it imposes on the bundling strategy of the products in an ad-hoc manner. On the other hand, the latter two models can make inference of preference interdependence directly based on the estimated parameters. These three models, however, consider only the choice situations that alternative options for each product categories are "buy" or "not buy". In the travel demand framework, there are at least two alternatives to be considered.

Hence, according to the above discussions, we extend certain aspects of the product bundling consideration, which were proposed by Manchanda *et al.* (1999) and Russell and Peterson (2000), so that these models can be applicable in two alternatives choice situation. In the next section, two discrete choice models considering interrelated recreational travel demand are formulated.

### 3. MODELLING APPROACH

#### 3.1 Specification of the Travelers' Choice Situation

Let  $n$  denote a traveler who faces a choice situation, in which he chooses one of two alternative travel modes and one of two alternative destinations. We define this situation as two-categories simultaneous choice behavior. The choice situation is illustrated in Figure 1.

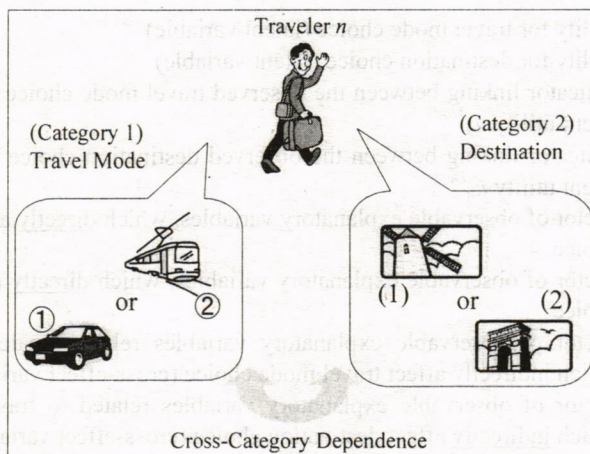


Figure 1. Specification of Choice Situation

Moreover, all variables appearing in both travel mode choice and destination choice are expressed as the difference of each two alternatives (e.g. the attribute of alternative No.1 minus the attribute of alternative No.2).

Under these assumptions, first, the formulation of a bivariate probit model is reported in section 3.2. Second, the formulation with multivariate logistic model (multinomial logit model) is shown in section 3.3. Although the formulations of both models are based on random utility maximization, the structures of the models are quite different and independent.

### 3.2 Model 1: Bivariate Dichotomous Probit Model with Cross-Category Dependence

In model 1, we represent the interaction between travel mode choice and destination choice by introducing these two elements such as:

- (i) Variables which affect another choice category alternately (Cross-Category Variables),
- (ii) Correlation structure between the utility, which constructs travel mode choice, and the utility that constructs destination choice.

Each travel choice behavior is represented mathematically as follows. These two choice models are now formulated separately, but are integrated by their correlation structure later.

#### Travel Mode Choice Model

$$u_n^{Mode} = \beta^{Mode} \mathbf{x}_n^{Mode} + \gamma^{Dest} \mathbf{z}_n^{Dest} + \varepsilon_n^{Mode} \quad (1)$$

$$d_n^{Mode} = \begin{cases} 1 & \text{if } u_n^{Mode} > 0 : \text{the case that Travel Mode 1 is chosen} \\ 0 & \text{if } u_n^{Mode} \leq 0 : \text{the case that Travel Mode 2 is chosen} \end{cases} \quad (2)$$

#### Destination Choice Model

$$u_n^{Dest} = \beta^{Dest} \mathbf{x}_n^{Dest} + \gamma^{Mode} \mathbf{z}_n^{Mode} + \varepsilon_n^{Dest} \quad (3)$$

$$d_n^{Dest} = \begin{cases} 1 & \text{if } u_n^{Dest} > 0 : \text{the case that Destination 1 is chosen} \\ 0 & \text{if } u_n^{Dest} \leq 0 : \text{the case that Destination 2 is chosen} \end{cases} \quad (4)$$

Where,

- $u_n^{Mode}$  : utility for travel mode choice (latent variable)
- $u_n^{Dest}$  : utility for destination choice (latent variable)
- $d_n^{Mode}$  : indicator linking between the observed travel mode choice behavior and the latent utility  $u_n^{Mode}$
- $d_n^{Dest}$  : indicator linking between the observed destination choice behavior and the latent utility  $u_n^{Dest}$
- $\mathbf{x}_n^{Mode}$  : vector of observable explanatory variables, which directly affect travel mode choice
- $\mathbf{x}_n^{Dest}$  : vector of observable explanatory variables, which directly affect destination choice
- $\mathbf{z}_n^{Dest}$  : vector of observable explanatory variables related to destination choice, which indirectly affect travel mode choice (cross-effect variables)
- $\mathbf{z}_n^{Mode}$  : vector of observable explanatory variables related to travel mode choice, which indirectly affect destination choice (cross-effect variables)
- $\varepsilon_n^{Mode}, \varepsilon_n^{Dest}$  : unobserved disturbances, which follow bivariate normal distribution. Here the mean of  $\varepsilon_n^{Mode}$  and  $\varepsilon_n^{Dest}$  are both 0's, the variances of them are both 1's, and their covariance is  $\rho$  (unknown parameter)

$\beta^{Mode}, \gamma^{Dest}, \beta^{Dest}, \gamma^{Mode}$  : vectors of unknown parameters

Among these variables, the cross-effect variables:  $z_n^{Mode}$  and  $z_n^{Dest}$  seem to be difficult to understand. Hence, we show one concrete example in order to understand the meaning of these values at first. Here, let us specify these variables as:

- $z_n^{Dest}$  : index of attractiveness for traveler  $n$  to have for the recreational destinations
- $z_n^{Mode}$  : index of comfortableness for traveler  $n$  to have for the travel modes

The coefficients for the cross-effects variables:  $\gamma^{Mode}$  and  $\gamma^{Dest}$  represent the magnitude of cross-effect variables:  $z_n^{Mode}$  and  $z_n^{Dest}$  respectively. For example, if the sign of the estimated parameter  $\gamma^{Dest}$  is positive, it means that an increase in the attractiveness of Destination (1) brings about an increase in utility of Travel Mode ①. In this case, the attractiveness of Destination (1) and the utility of Travel Mode ① are positive covariates. On the other hand, if the sign of the estimated parameter  $\gamma^{Mode}$  is negative, it follows that an increase in the comfortableness of Travel Mode ① brings about an increase in utility of Destination (2). In this way, the estimated parameters of  $\gamma^{Mode}$  and  $\gamma^{Dest}$  directly represent the cross-category dependence between travel mode choice and destination choice: whether they are the complements or substitutes of each other. This insight is very meaningful for understanding the effects of the policy for travel mode on the recreational destinations.

Next, the correlated error structure of  $(\varepsilon_n^{Mode}, \varepsilon_n^{Dest})$  represents the co-incidence relationship between travel mode ① or ② and destination (1) or (2). We can understand the specific relationship of co-incidence by estimated parameter  $\rho$ . For example, if  $\rho > 0$ , then an increase in the travel mode ①'s utility or a decrease in the travel mode ②'s utility will lead to an increase in the utility of destination (1) relatively. In other words, the error covariance  $\rho$  captures the linkage between the uncontrollable and unobservable factors that drive simultaneous choices of travel mode and destination. Of course, the error structure arises because of the model misspecification (such as omitted variables or not considering the travelers' choice sets). This must be kept in mind while drawing statistical inferences based on the error correlation terms.

Finally, the above model formulation of the simultaneous choice of travel mode and destination results in a bivariate probit model. It is important to note that the bivariate probit model is quite different from the *binomial* probit model or *multinomial* probit model. The latter two models only allow the choice of one alternative from a set of mutually exclusive two or more alternatives. Finally, the simultaneous choice probability of travel mode and destination indexed by  $d_n^{Mode}$  and  $d_n^{Dest}$  is formulated as:

$$\Pr(d_n^{Mode}, d_n^{Dest}) = \Phi_2(w_n^{Mode}, w_n^{Dest}, \rho^*) \tag{5}$$

where  $\Phi_2(\cdot)$  denotes the bivariate cumulative normal distribution and the definitions of other functions are as follows. For example, the equation (10) denotes the bivariate normal distribution and the equation (9) denotes the cumulative distribution of it.

$$w_n^{Mode} = (2d_n^{Mode} - 1) (\beta^{Mode} x_n^{Mode} + \gamma^{Dest} z_n^{Dest}) \tag{6}$$

$$w_n^{Dest} = (2d_n^{Dest} - 1) (\beta^{Dest} x_n^{Dest} + \gamma^{Mode} z_n^{Mode}) \tag{7}$$

$$\rho^* = (2d_n^{Mode} - 1)(2d_n^{Dest} - 1)\rho \tag{8}$$

$$\Phi_2(s_1, s_2, \sigma) = \int_{-\infty}^{s_1} \int_{-\infty}^{s_2} \phi_2(q_1, q_2, \sigma) dq_1 dq_2 \tag{9}$$

$$\phi_2(q_1, q_2, \sigma) = \frac{1}{2\pi(1-\sigma^2)^{1/2}} \exp\left(\frac{q_1^2 + q_2^2 - 2\sigma q_1 q_2}{2\sigma^2 - 2}\right) \tag{10}$$

We can estimate  $\beta^{Mode}$ ,  $\gamma^{Dest}$ ,  $\beta^{Dest}$ ,  $\gamma^{Mode}$  and  $\rho$  by applying the maximum likelihood method to the formulation (5). The parameters of the bivariate probit models (including the covariance parameter  $\rho$ ) can be easily computable by using econometric software such as LIMDEP (Econometric Software Inc., 1998).

### 3.3 Model 2: Bivariate Dichotomous Logit Model with Cross-Category Dependence

It is the best way to induce the joint (simultaneous) choice probability that a travel mode and a recreational destination are chosen together. However, in real travel behavior situation, it is difficult to estimate the joint probability directly, since it requires a detailed understanding of cross-category demand relationships (interdependence of different choice categories). The modeling approach proposed here is in accordance with Russell and Peterson (2000), and can avoid these problems.

Suppose that a traveler  $n$  has made choices in recreational destination and is now considering which travel mode is to be selected. The proposed approach assumes that we can specify  $p(\text{Travel Mode} | \text{Destination})$ , the probability of selecting the particular travel mode, given the known outcomes of the destination choice. In the same way,  $p(\text{Destination} | \text{Travel Mode})$  can be observed. The probability  $p(\text{Travel Mode} | \text{Destination})$  and  $p(\text{Destination} | \text{Travel Mode})$  are called "full conditional distributions" (Russell and Peterson, 2000). By using the principles of Markov Random Field, we can induce the joint probability  $p(\text{Travel Mode}, \text{Destination})$ .

In model 2, we represent the interaction between travel mode choice and destination choice by introducing these two elements such as:

- (i) variables which affects another choice category alternately (Cross-Category Variables); the same as that of Model 1,
- (ii) conditional utility function for particular choice category, which includes the choice results of other category as dummy variable.

The conditional utility (or choice probability) method requires the researchers to specify the probability that each choice category such as travel mode and destination will be given, conditioned that the choice results of all other choice categories are known. For example, let us define the conditional utility of traveler  $n$  for travel mode conditional upon the result of destination choice as:

#### (Conditional) Travel Mode Choice Model

$$u_n^{Mode} = \beta^{Mode} x_n^{Mode} + \gamma^{Dest} z_n^{Dest} + \theta_{Dest}^{Mode} d_n^{Dest} + v_n^{Mode} \quad (11)$$

$$d_n^{Mode} = \begin{cases} 1 & \text{if } u_n^{Mode} > 0 : \text{the case that Travel Mode 1 is chosen} \\ 0 & \text{if } u_n^{Mode} \leq 0 : \text{the case that Travel Mode 2 is chosen} \end{cases} \quad (12)$$

In the same way, the conditional utility of traveler  $n$  for destination conditional upon the result of travel mode choice as:

#### (Conditional) Destination Choice Model

$$u_n^{Dest} = \beta^{Dest} x_n^{Dest} + \gamma^{Mode} z_n^{Mode} + \theta_{Mode}^{Dest} d_n^{Mode} + v_n^{Dest} \quad (13)$$

$$d_n^{Dest} = \begin{cases} 1 & \text{if } u_n^{Dest} > 0 : \text{the case that Destination 1 is chosen} \\ 0 & \text{if } u_n^{Dest} \leq 0 : \text{the case that Destination 2 is chosen} \end{cases} \quad (14)$$

where the definitions of  $u_n^{Mode}$ ,  $u_n^{Dest}$ ,  $d_n^{Mode}$ ,  $d_n^{Dest}$ ,  $x_n^{Mode}$ ,  $x_n^{Dest}$ ,  $z_n^{Mode}$ ,  $z_n^{Dest}$ ,  $\beta^{Mode}$ ,  $\gamma^{Dest}$ ,  $\beta^{Dest}$  and  $\gamma^{Mode}$  are the same as those in Model 1. We assume that the value of  $d_n^{Dest}$  is already known in the equation (11), whereas the value of  $d_n^{Mode}$  is known in the equation (13). The definitions of other variables are as follows:

- $v_n^{Mode}, v_n^{Dest}$  : the unobserved disturbances, which follow independent and identically logistic distribution
- $\theta_{Dest}^{Mode}, \theta_{Mode}^{Dest}$  : the unknown parameters, which specify the interdependence of, travel mode choice and destination choice

Note that  $\theta_{Dest}^{Mode} > 0$  implies a positive association between the choice of travel mode ① and the choice of destination (1), while  $\theta_{Dest}^{Mode} < 0$  implies a negative association. In that sense,  $\theta_{Dest}^{Mode}$  and  $\theta_{Mode}^{Dest}$  are the same indicators as error-covariance parameter  $\rho$  shown in Model 1.

Finally, the conditional probability of selecting travel mode ① can be express as the binomial logit model conditioned that the choice result of destination  $d_n^{Dest}$  is known.

$$\begin{aligned} \Pr(d_n^{Mode} = 1 \mid d_n^{Dest}) &= \Pr(u_n^{Mode} > 0 \mid d_n^{Dest}) \\ &= \frac{1}{1 + \exp\left[-\left(\beta^{Mode} \mathbf{x}_n^{Mode} + \gamma^{Dest} \mathbf{z}_n^{Dest} + \theta_{Dest}^{Mode} d_n^{Dest}\right)\right]} \end{aligned} \quad (15)$$

In the same way, the conditional probability of selecting destination (1) can be express as:

$$\begin{aligned} \Pr(d_n^{Dest} = 1 \mid d_n^{Mode}) &= \Pr(u_n^{Dest} > 0 \mid d_n^{Mode}) \\ &= \frac{1}{1 + \exp\left[-\left(\beta^{Dest} \mathbf{x}_n^{Dest} + \gamma^{Mode} \mathbf{z}_n^{Mode} + \theta_{Mode}^{Dest} d_n^{Mode}\right)\right]} \end{aligned} \quad (16)$$

The equation (15) and equation (16) are the full conditional models in the context of travel mode choice and destination choice. However, they are just the “conditional” distributions and we need to formulate the joint distribution of the simultaneous choices of travel mode and destination. For this task, we employ the “Factorization Theorem” proposed by Besag (1974) in accordance with Russell and Peterson (2000). The proof of the theorem is detailed in Cressie (1993). The theorem provides a simple mathematical way of deriving a joint distribution given a set of full conditional distributions. To adopt the theorem to our model specification, the symmetry of the interdependence parameters should be imposed as:

$$\theta_{Dest}^{Mode} = \theta_{Mode}^{Dest} = \theta \quad (17)$$

Under this assumption, the joint probability that, travel mode choice  $d_n^{Mode}$  takes a value of  $a$  and destination choice  $d_n^{Dest}$  takes value of  $b$ , is as follows:

$$\Pr(d_n^{Dest} = a, d_n^{Mode} = b) = \frac{\exp(V_n(a, b))}{\sum_{a' \in \{0,1\}} \sum_{b' \in \{0,1\}} \exp(V_n(a', b'))} \quad (18)$$

where  $a$  and  $b$  are the dummy variables which take the values 1 or 0 dependent on the travelers' choice. And then,  $V_n(a, b)$  is defined as:

$$V_n(a, b) = a \left( \beta^{Mode} \mathbf{x}_n^{Mode} + \gamma^{Dest} \mathbf{z}_n^{Dest} \right) + b \left( \beta^{Dest} \mathbf{x}_n^{Dest} + \gamma^{Mode} \mathbf{z}_n^{Mode} \right) + ab\theta \quad (19)$$

Although the model formulations (18) and (19) are derived as the logical conclusion of applying the factorization theorem to equations (15) and (16), the emerged formulations (18) and (19) are equivalent to the multinomial logit formulation whose deterministic part of the utility function are defined as (19) and whose choice sets are the  $2^2$  combinations of travel mode and destinations. The deterministic part of the global utility function  $V_n(a, b)$  specifies the simultaneous choice of travel mode and destination indexed by the pair of  $(a, b)$ . Hence, the unknown parameters  $\beta^{Mode}$ ,  $\gamma^{Dest}$ ,  $\beta^{Dest}$ ,  $\gamma^{Mode}$  and  $\theta$  are estimated by usual multinomial logit estimation procedure. The interdependence parameter  $\theta$  and the cross-category

dependence parameters  $\gamma^{Dest}$ ,  $\gamma^{Mode}$  specify the interactions between travel mode choice and destination choice of travelers.

#### 4. EMPIRICAL APPLICATION

Using individuals travel survey data conducted in Japan, the two models proposed in previous section are verified. We present a brief description of the data, the specification and operationalization of explanatory variables, and estimation results. We then discuss the results of the proposed model and compare them. We conclude this chapter with the brief description on model prediction.

##### 4.1 Details of the Survey Data and the Definition of Choice Alternatives

The data were taken from recreational travel survey conducted by Ministry of Construction in Japan in 1997. (Refer to Mizokami and Furu-ichi, 1998, for the detail of the survey). The survey was executed for examining the personal and travel characteristics of travelers who visited Nara Prefecture in Japan. It is well known that Nara Pref. is one of the most famous recreational regions in Japan and has the many recreational resources such as old shrines and temples. Various information on travelers recreational trips to Nara Pref. were surveyed.

As we mentioned above, we model the simultaneous choice behavior of travelers' travel mode choice and recreational destination choice as shown in Figure 1. As for the travel mode, the privately-owned car is labeled as ①, whereas the railway is labeled as ②. These are the choice results of travelers who choose them as the main transportation means.

On the other hand, we need to divide the region into two destination zones. The criteria we adopted for zoning are: the regional differences of these two destinations are defined as (1): Urban Area which includes Nara-City, Ikoma-City and Yamato-Koriyama-City, and (2): Suburbs Area the north half part of Nara Prefecture. There exist many famous historical temples and architectures in the former area, whereas the latter includes many natural recreational resources such as scenic spots, outdoor sports places and so on. Hence, we think those areas are well differentiated. The zonings of these destinations are shown in Figure 2.

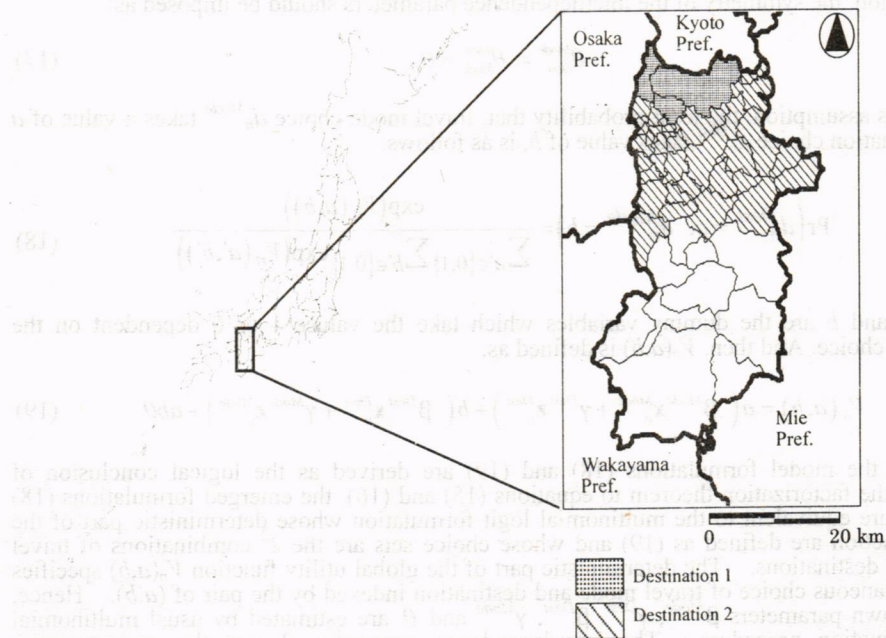


Figure 2. Zoning of Destination Areas (Nara Prefecture in Japan)



Under these assumptions on choice alternatives, we focus on the simultaneous choice behavior on the pair of travel mode and destination. Here, the recreational travel data on the first visited area of the travelers living in all of the Kansai-Area and the part of the Chubu-Area. We extracted samples with full response to the variables of our interest (see Table 2). Finally, the samples of total 1,243 travelers' data were used for the identification of the models. The cross tabulation of the choice results of these travelers are shown in Table 1.

Table 1. Cross Tabulation of Choice Results (Travel Mode vs. Destination)

Travel Mode Destination	Mode ① (Car)	Mode ② (Railway)	Total
Destination 1 (Urban Area)	425	145	570
Destination 2 (Suburbs Area)	501	172	673
Total	926	317	1,243

#### 4.2 Specifications of Explanatory Variables

To specialize the cross-category dependence choice models as shown in previous section, we need to define the explanatory variables. As we mentioned before, all variables are expressed as the difference of two alternatives. The outlines of variable definitions are displayed in Table 2.

As for the variable: *Gender*, *Age*, *Accomp*, *NVehicle*, *Log(Rec)* and *Log(Inc)*, we can obtain them from the travel survey data directly. On the other hand, the level of service (LOS) variables (*MTime*, *MCost*, *DTime* and *DCost*) and the subjective ratings (*Att1*, *Att2*, *Comf1* and *Comf2*) cannot be observed except for the data on the choice pairs of travel mode and destination the traveler actually chose. Hence, these data were substituted by the mean values over the observed samples. Under these assumptions, the variables such as *MTime*, *MCost*, *DTime*, *DCost*, *Comf* and *Attract* are defined as shown in Table 2.

As for the cross-effect variable  $z_n^{Dest}$ , three variables (*DTime*, *Att1* and *Att2*) are adopted. By employing *DTime* as the cross-effect variable, we can infer how the transportation improvement within each zonal level affects travel mode choice behavior. In the Attractiveness ratings of Destination 1 and Destination 2, the categorization of the rating ranges from best (5) to worst (1). In the same way, *MTime* and the subjective rating for the comfortableness of the private car (*Comf1*) and the railway (*Comf2*) are used for the cross-effect variable  $z_n^{Mode}$ .

Table 2. Definitions of Explanatory Variables

(a) Definition of $x_n^{Mode}$ (which can be seen in Eq.(1) and Eq.(11): Travel Mode Choice Model)	
<i>MTime</i>	The mean travel time of destination 1 and destination 2 from home (Minute) (calculated for every travel mode, and differenced value is employed)
<i>MCost</i>	The mean travel cost per person of destination 1 and destination 2 from home (Yen) (calculated for every travel mode, and differenced value is employed)
<i>Age</i>	Age of the respondent
<i>Gender</i>	1 if the respondent is male, 0 if female.
<i>Accomp</i>	1 if the respondent is accompanied by someone, 0 otherwise
<i>NVehicle</i>	Number of private cars owned by each household
<i>Log(Rec)</i>	Logarithm of the sightseeing expenditure for the respondent per year (log(Yen))
<i>Comf</i>	Comfortableness ratings of each travel mode (5-point scale; 1:worst – 5:best) (= <i>comf1</i> – <i>comf2</i> : calculated for every travel mode, and differenced value is employed)
(b) Definition of $z_n^{Dest}$ (which can be seen in Eq.(1) and Eq.(11): Travel Mode Choice Model)	
<i>DTime</i>	The mean travel time of travel mode 1 and travel mode 2 from home (Minute) (calculated for every destination and differenced value is employed)
<i>Att1</i>	Attractiveness ratings of Destination 1 (5-point scale; 1:worst – 5:best)
<i>Att2</i>	Attractiveness ratings of Destination 2 (5-point scale; 1:worst + 5:best)
(c) Definition of $x_n^{Dest}$ (which can be seen in Eq.(2) and Eq.(12): Destination Choice Model)	
<i>DTime</i>	The mean travel time of travel mode 1 and travel mode 2 from home (Minute) (calculated for every destination and differenced value is employed)
<i>DCost</i>	The mean travel cost per person of travel mode 1 and travel mode 2 from home (Yen) (calculated for every destination and differenced value is employed)
<i>Age</i>	Age of the respondent
<i>Gender</i>	1 if the respondent is male, 0 if female.
<i>Accomp</i>	1 if the respondent is accompanied by someone, 0 otherwise
<i>Log(Inc)</i>	Logarithm of the income for the respondent per year (log(Yen))
<i>Attract</i>	Attractiveness ratings of each destination (5-point scale; 1:worst – 5:best) (= <i>Att1</i> – <i>Att2</i> : calculated for every travel mode, and differenced value is employed)
(d) Definition of $z_n^{Mode}$ (which can be seen in Eq.(2) and Eq.(12): Destination Choice Model)	
<i>MTime</i>	The mean travel time of destination 1 and destination 2 from home (Minute) (calculated for every travel mode, and differenced value is employed)
<i>Comf1</i>	Comfortableness ratings of Travel Mode ① (5-point scale; 1:worst – 5:best)
<i>Comf2</i>	Comfortableness ratings of Travel Mode ② (5-point scale; 1:worst – 5:best)

### 4.3 Estimation Results

The estimation results are shown in Table 3 and Table 4. Table 3 presents the estimation results of Model 1 (Cross-Category Probit) and Table 4 for Model 2 (Cross-Category Logit). On the left side of each table, the estimation results of the *Null* model can be seen, whereas that of the *Full* model on the right side. The *Null* model in both Table 3 and Table 4 ignores the cross-category dependence between travel mode choice and destination choice by discarding the variables  $z_n^{Dest}$  and  $z_n^{Mode}$ . In each model (Model 1 and Model 2), It is obvious that the goodness of fit of the *Full* Model is greater than that of the *Null* Model, judging from the value of McFadden's adjusted likelihood ratio.

Most of the estimates come out with the expected sign and are significantly different from zero at the 5% significance level, except for several parameters such as *constants*, *Log(Rec)*,

*Log(Inc)* and *Attract*. Each parameter has the same sign in both Model 1 and Model 2, though some have different absolutes. The differences in the absolutes of the parameters may be a result of the change in the scales of the models. On the other hand, the parameters related to Level of Services such as *MTime*, *MCost*, *DTime* and *DCost* take positive signs. It implies that travelers do not always regard the travel time or cost as the traveling resistance for recreational trips. It is mainly because that, for travelers, the din and bustle can be regarded as an attractiveness of recreational sites to some extent. This is one of the specific characteristics of recreational travel behavior (e.g. International Association of Traffic and Safety Sciences, 1998).

We find that the cross-effect variables such as *Att1* and *Att2* in  $z^{Dest}$  take negative coefficients but not statistically significant. It means that the use of Travel Mode ② (Railway) is promoted if the attractiveness specific to each destination increases. On the other hand, the coefficients of other cross-effect variable: *Comf1* is negative, whereas that of *Comf2* is positive. It follows that travelers visit Destination (1) more frequently than Destination (2) if the comfortableness specific to Travel Mode ② increases. Moreover, the magnitudes of the coefficients of cross-effects parameters *DTime* are almost the same as the parameters *DTime* included in the own-effects in both Table 3 and Table 4. This is true of the parameter *MTime* too. These findings imply that the influence of the cross effects are as strong as their own-effects.

On the other hand, the estimates such as  $\rho$  in Model 1 and  $\theta$  in Model 2 show the co-occurrence pattern of the particular pair of travel modes and destinations. In both models, these co-occurrence indicator parameters are in the negative sign. This implies that an relative increase in the utility of the private car will lead to an increase in the share of the travelers who visit suburban areas. Judging from both the parameter  $\rho$  and  $\theta$  and the McFadden's adjusted likelihood index, the estimation results of Model 1 show the fitness as good as those of Model 2. Of course, these results do not imply the general properties of the recreational trips, because this empirical study is just a specific experiment. The more empirical cases should be analyzed to generalize the results.

Table 3. Estimation Results of Model 1 (Cross-Category Probit): Comparison of the Null Model and the Full Model

The Null Model for Model 1 (discarding the variables $z$ )			The Full Model for Model 1		
Part of Travel Mode Choice		Part of Destination Choice	Part of Travel Mode Choice		Part of Destination Choice
Variables	Estimates ( $t$ -static)	Variables	Estimates ( $t$ -static)	Variables	Estimates ( $t$ -static)
$(x^{Mode})$		$(x^{Dest})$		$(x^{Dest})$	
$MTime \times 10^{-2}$	0.632 (2.86)	$DTime \times 10^{-2}$	0.344 (4.14)	$MTime \times 10^{-2}$	0.586 (6.13)
$MCost \times 10^{-5}$	10.2 (3.65)	$DCost \times 10^{-5}$	16.7 (4.25)	$DCost \times 10^{-5}$	7.26 (2.41)
$Age \times 10^{-2}$	-1.91 (-6.17)	$Age \times 10^{-2}$	-1.72 (-6.17)	$Age \times 10^{-2}$	-1.66 (-5.24)
$Gender$	0.369 (4.18)	$Gender$	0.173 (1.86)	$Gender$	0.337 (3.80)
$Accomp$	0.914 (5.16)	$Accomp$	0.542 (2.85)	$Accomp$	0.817 (4.30)
$NVehicle$	0.329 (8.53)	$NVehicle$	3.65 (0.890)	$NVehicle$	0.306 (7.51)
$Compf$	0.139 (4.07)	$Log(Inc) \times 10^{-2}$	-3.07 (-1.57)	$Compf$	0.162 (4.61)
$Log(Rec) \times 10^{-2}$	4.81 (1.38)			$Log(Inc) \times 10^{-2}$	3.97 (1.08)
$(z^{Mode})$		$(z^{Mode})$		$(z^{Mode})$	
				$MTime \times 10^{-2}$	0.567 (6.00)
				$Compf1 \times 10^{-2}$	-9.18 (-1.49)
				$Compf2 \times 10^{-2}$	-6.74 (-1.04)
Const.	-0.131 (-0.52)	Const.	0.260 (1.09)	Const.	0.635 (1.37)
Parameter for Covariance of the Error Term $\rho = -0.117(-2.10)$					
$L(0) = -1723.2, L_{Max} = -1415.1, McFadden's \text{ adjusted likelihood ratio} = 0.175$					

Table 4. Estimation Results of Model 2 (Cross-Category Logit): Comparison of the Null Model and the Full Model

The Null Model for Model 2 (discarding the variables $z$ )			The Full Model for Model 2		
Part of Travel Mode Choice		Part of Destination Choice	Part of Travel Mode Choice		Part of Destination Choice
Variables	Estimates ( $t$ -static)	Variables	Estimates ( $t$ -static)	Variables	Estimates ( $t$ -static)
$(x^{Mode})$		$(x^{Dest})$		$(x^{Dest})$	
$MTime \times 10^{-2}$	0.111 (1.16)	$DTime \times 10^{-2}$	0.629 (3.00)	$MTime \times 10^{-2}$	1.18 (5.64)
$MCost \times 10^{-5}$	18.6 (3.71)	$DCost \times 10^{-5}$	25.5 (4.74)	$DCost \times 10^{-5}$	12.4 (2.40)
$Age \times 10^{-2}$	-3.64 (-6.69)	$Age \times 10^{-2}$	-3.01 (-6.40)	$Age \times 10^{-2}$	-3.27 (-5.89)
$Gender$	0.662 (4.35)	$Gender$	0.312 (2.02)	$Gender$	0.611 (3.94)
$Accomp$	1.59 (5.33)	$Accomp$	0.973 (3.15)	$Accomp$	1.46 (4.77)
$NVehicle$	0.687 (7.10)	$NVehicle$	5.92 (0.894)	$NVehicle$	0.623 (6.36)
$Compf$	0.264 (4.70)	$Log(Inc) \times 10^{-2}$	-4.76 (-1.51)	$Compf$	0.281 (4.87)
$Log(Rec) \times 10^{-2}$	8.09 (1.27)			$Log(Inc) \times 10^{-2}$	6.92 (1.07)
$(z^{Mode})$		$(z^{Mode})$		$(z^{Mode})$	
				$MTime \times 10^{-2}$	1.14 (5.48)
				$Compf1 \times 10^{-2}$	-17.1 (-1.38)
				$Compf2 \times 10^{-2}$	-11.9 (-1.08)
Const.	-0.113 (-0.259)	Const.	0.146 (0.291)	Const.	1.34 (1.64)
Cross-Category Dependence Parameter $\theta = -0.295 (-2.09)$					
$L(0) = -1723.2, L_{Max} = -1409.9, McFadden's \text{ adjusted likelihood ratio} = 0.177$					
Cross-Category Dependence Parameter $\theta = -0.417 (-2.80)$					
$L(0) = -1723.2, L_{Max} = -1374.8, McFadden's \text{ adjusted likelihood ratio} = 0.197$					

#### 4.4 Comparison of Elasticity Analysis by Cross-Category Probit Model

The estimated parameters of choice models in preceding sections can be misleading. To be concrete, the absolute scale of the parameters gives a distorted picture of the response of the dependent variable to a change in one of the stimuli, since the models are actually of a series of probability. Moreover, from a managerial perspective on transport facilities or recreational sites, the most interesting aspects of this research are to execute the elasticity analysis in order to verify the influence of the structures of the proposed models on the demand forecasting.

In this section, we compute the elasticity of the choice share of the private cars with respect to a particular explanatory variables. Here, we especially focus on only Model 1 to compare the detailed results between some null models and the full model. The outlines of these models are as follows:

- The Null Model 1: the model that accounts for neither cross-category variables  $\mathbf{z}$  nor the correlation of error terms between travel mode choice and destination choice (equivalent to fitting a separate probit regression for each category).
- The Null Model 2: the model that accounts for cross-category variables  $\mathbf{z}$  but still do not account for the error correlation (equivalent to the Null model of Model 1 shown in chapter 3).
- The Full Model: the bivariate probit model which incorporates both cross-category variables and the error correlation.

When any of the covariates  $y$  increases by 1% from initial value, the marginal elasticity for bivariate probit model is calculated by the following formula. The computation procedure is detailed in Greene (1996, 1997).

$$\begin{aligned} & \frac{\partial E[d^{Mode} | \mathbf{x}^{Mode}, \mathbf{x}^{Dest}, \mathbf{z}^{Mode}, \mathbf{z}^{Dest}, d^{Dest} = 1]}{\partial y} \\ &= \frac{\partial}{\partial y} \left( \frac{\Phi_2(w^{Mode}, w^{Dest}, \rho^*)}{\Phi(w^{Mode})} \right) \\ &= \left( \frac{g_{Mode}(w^{Mode}, w^{Dest}, \rho^*)}{\Phi(w^{Dest})} \right) \beta_{Mode, y} \\ &+ \left( \frac{\Phi(w^{Dest}) g_{Dest}(w^{Mode}, w^{Dest}, \rho^*) - \Phi_2(w^{Mode}, w^{Dest}, \rho^*) \phi(w^{Dest})}{[\Phi(w^{Dest})]^2} \right) \beta_{Dest, y} \quad (20) \end{aligned}$$

Where each function and variable are defined as follows.

$$g_{Mode}(w_{Mode}, w_{Dest}, \rho^*) \equiv \frac{\Phi_2(w_{Mode}, w_{Dest}, \rho^*)}{\partial w_{Mode}} = \phi(w_{Mode}) \Phi \left( \frac{w_{Dest} - \rho^* w_{Mode}}{\sqrt{1 - \rho^{*2}}} \right) \quad (21)$$

(and likewise for  $g_{Dest}(w_{Mode}, w_{Dest}, \rho^*)$ )

- $y$ : any of the explanatory variable among  $\mathbf{x}_n^{Mode}$ ,  $\mathbf{x}_n^{Dest}$ ,  $\mathbf{z}_n^{Dest}$  and  $\mathbf{z}_n^{Mode}$
- $\beta_{Mode, y}$ ,  $\beta_{Dest, y}$ : the coefficient on the target variable  $y$ .
- $\phi(\bullet)$ : the density function of the univariate standard normal distribution.
- $\Phi(\bullet)$ : the univariate standard cumulative normal distribution function

In above formulations, the suffix  $n$ , which denotes the traveler, is eliminated.

Each elasticity value for the above three models are respectively given by the formula (20) and computed with the parameter estimates and some configuration of the data. We compute the marginal elasticities at the sample means of the variables. The sample means shown in Table 5. The values reported are the percentage changes. That is, the table shows the values of  $100/E[\dots] \times \partial E[a^{Mode} | a^{Dest}, y] / \partial y$ .

For each of the three models, the computed elasticity values are shown in Table 6. The elasticity results employed by only three variables are shown for lack of space. The sensitivity of each variable of the full model is largest among these three models. This means that the existence of cross-category dependence such as variables and error correlation promotes the sensitivity of travel demands.

Table 5. Mean Values of Explanatory Variables (employed for Elasticity Analysis)

Variables	Mean	Variables	Mean
<i>MTime</i>	16.32	<i>Log(Rec)</i>	3.018
<i>MCost</i>	597.2	<i>Log(Inc)</i>	5.054
<i>DTime</i>	-18.91	<i>Att1</i>	3.997
<i>DCost</i>	-237.3	<i>Att2</i>	1.521
<i>Age</i>	42.95	<i>Att</i>	-0.002170
<i>Gender</i>	0.5873	<i>Comf1</i>	3.145
<i>Accomp</i>	0.9496	<i>Comf2</i>	3.649
<i>NVehicle</i>	1.422	<i>Comf</i>	-0.4994

Table 6. Results of Elasticity Analysis (percentage changes)

	Variables to be increased by 1%		
	<i>Acomp</i>	<i>NVehicle</i>	<i>Comf</i>
Null Model 1	0.2759	0.1003	0.0442
Null Model 2	0.3104	0.1068	0.1003
Full Model	0.2885	0.1012	0.0537

## 5. CONCLUSION AND FUTURE PROSPECTS

In this paper, we proposed a new concept on interdependence behaviors within the recreational trips and a framework of recreational travel behavior model that considers the interrelations among different choices, such as destination choice and travel mode choice. In order to specify the interrelations such as "cross-category dependence" or "co-incidence correlation", the concept of "product bundling" or "Market Basket Analysis" is introduced to the discrete choice model according to the marketing science literature. In this context, we attempted to formulate two models. The first model is based on multivariate probit formulation and the second model is based on multivariate logistic distribution formulation or multinomial logit formulation. This framework could be useful in formulating the marketing strategies for recreational sites, because it can make suggestions about what kind of combinations of recreational components are more attractive to consumers. In order to verify the models empirically, the behavioral survey data extracted from the recreational travel survey in Nara-Prefecture, Japan, are employed. The empirical results confirmed that there exist significant interdependences between travel mode choice and destination choice, and that the proposed models are superior to the Null models in terms of model fitness. Our analysis also revealed that the proposed framework may sometimes bring about drastic change in policy simulation compared with the Null models.

However, there is a great deal of scope for further works. First, one of the most important works is to continue more empirical studies based on various recreational trip data, since the most of the findings derived from this research did not confirm their generalities. Especially, independent surveys could be considered so as to validate the model results. Second, the interdependence between two choice categories should be more specified by other than the cross-category variables. One of the promising approaches is to adopt the recursive simultaneous equations model involving two binary choice variables (Wilde, 2000). With this approach, the choice result of another category will be specifically incorporated into the models as choice dummy variables. Third, the proposed models cannot deal with the choice situations that have more than three alternatives to be considered in one category. For example, if there exist three travel modes such as car, bus and railway, both of the proposed two models are not applicable. In order to solve this problem, the multinomial random variables should be incorporated (Strauss, 1977). Finally, for better understanding of bundling phenomenon or cross-category consideration, we need to look back on the principles of human decision-making process. How human regards the several items as one category is one of the most discussed topics in psychological literature, called "Categorical Perception" (e.g. Harnad, 1987). We strongly believe that such psychological phenomenon research can be appropriate for understanding the mechanism of the private and not-constrained behavior such as the recreational trips.

## REFERENCES

### a) Books and Books chapters

- Cressie, N. (1993) **Statistics for Spatial Data**, John Wiley and Sons, New York.
- Greene, W. (1997) **Econometric Analysis**, Prentice-Hall, New Jersey.
- Harnad, S. (1987) **Categorical perception : the groundwork of cognition**, The Cambridge University Press, Cambridge.
- International Association of Traffic and Safety Sciences (1998) **Attractive Recreational Spots and Transportation**. Gihodo Press, Tokyo (In Japanese).
- Koppelman, F. and Sethi, V. (2000) Closed-Form Discrete-Choice Models, In D. Hensher and K. Button (eds.), **Handbook of Transport Modelling**, Pergamon, Amsterdam.

### b) Journal papers

- Ben-Akiva, M. and Gershensfeld, S. (1998) Multi-Featured Products and Services: Analyzing Pricing and Bundling Strategies, **Journal of Forecasting**, Vol.17, pp.175-196.
- Besag, J. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems, **Journal of the Royal Statistical Society B**, Vol.36, pp.192-236.
- Chuang, J. and Sirbu, M. (1999) Optimal Bundling Strategy for Digital Information Goods: Network Delivery of Articles and Subscriptions, **Information Economics and Policy**, Vol.11, pp.147-176.
- Hanson, W. and Martin, R. (1990) Optimal Bundle Pricing, **Management Science**, Vol.36, pp.155-174.
- Manchanda, P., Ansari, A. and Gupta, S. (1999) The "Shopping Basket": A Model for Multicategory Purchase Incidence Decisions, **Marketing Science**, Vol.18, pp.95-114.
- Morikawa, T., Sasaki, K. and Azuma, R. (1995) Modeling Sightseeing Travel Behavior for Evaluation of Road Network Improvement in the Recreational Area, **Infrastructure Planning Review of JSCE**, No.12, pp.539-547 (in Japanese).
- Russell, G. and Petersen, A. (2000) Analysis of Cross Category Dependence in Market Basket Selection, **Journal of Retailing**, Vol.76, pp.367-392.

Russell, G., Bell, D., Bodapati, A., Brown, G., Chiang, J., Gaeth, G., Gupta, S. and Manchanda, P. (1997) Perspectives on Multiple Category Choice, **Marketing Letters**, Vol.8, pp.297-305.

Russell, G., Ratneshwar, S., Shocker, A., Bell, D., Bodapati, A., Degeratu, A., Hildebrandt, L., Kim, N., Ramaswari, S. and Shankar, V. (1999) Multiple-Category Decision-Making: Review and Synthesis, **Marketing Letters**, Vol.10, pp.319-332.

Strauss, D. (1977) Clustering on Coloured Lattices, **Journal of Applied Probability**, Vol.14, pp.135-143.

Tay, R., McCarthy, P. and Fletcher, J. (1996) A Portfolio Choice Model of the Demand for Recreational Trips, **Transportation Research B**, Vol.30, pp.325-337.

Train, K., McFadden, D. and Ben-Akiva, M. (1987) The Demand for Local Telephone Services: A Fully Discrete Choice Model of Residential Calling Patterns and Service Choices, **Rand Journal of Economics**, Vol.18, pp.109-123.

Wilde, J. (2000) Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors, **Economic Letters**, Vol.69, pp.309-312.

#### c) Papers presented to conferences

Mizokami, S. and Furuichi, E. (1998) Recreational Travel Survey in Nara Prefecture and its Outline, **Proceeding 34<sup>th</sup> Infrastructure Planning Symposium**, Tokyo, Japan, 25-26, November 1998 (in Japanese).

#### d) Other documents

Econometric Software Inc. (1998) **LIMDEP Version 7.0**, Econometric Software Inc., New York and Sydney.

Greene, W. (1996) Marginal Effects in the Bivariate Probit Model, **Working Paper**, EC-96-11, Stern School of Business, New York University, NY, USA.