The Need of Theoretical Identification of Latent-Class-Analysis Based Competitive-Market-Structure-Analysis Models

Akihiro INOUE*

The author introduces a system of non-spatial models of competitive-market-structure analysis, the MIGHT (Models to Identify competitive Groups with population Heterogeneity using confirmatory Tests) system, which attempts to test competitive-market-structure (CMS) hypotheses based on latent-class models for the analysis of brand-switching matrices. From the viewpoint of designing marketing strategy and plans, the identification of models of competitive-market-structure analyses is critical since the lack of identification can lead to incorrect designs. Model identification should be ensured *a priori and theoretically on the basis of the population*, rather than *post hoc* on the basis of a particular sample set. While prior to the MIGHT system no CMS model based on latent-class models secures the identification based on the population, the MIGHT system ensures *a priori*, theoretical identification. The MIGHT system consists of four submodels that are different with respect to the order of consumer stochastic choice behavior and heterogeneity in both choice behavior and consideration sets among submarkets. We compare the MIGHT system with the Grover and Srinivasan model and show the superiority of the MIGHT system in terms of the goodness-of-fit, the ease of interpretation, and the stability of parameter estimates. We also do a conceptual comparison of the MIGHT system with other CMS models and demonstrate its greater flexibility and less restrictiveness.

1. Introduction

In this paper we present a system of non-spatial models of competitive-market-structure analysis, the MIGHT (Models to Identify competitive Groups with population Heterogeneity using confirmatory Tests) system, which attempts to test CMS hypotheses based on latent-class models for the analysis of brand-switching matrices and assuring the identification *a priori and theoretically on the basis of the population*, rather than *post hoc* on the basis of a particular sample-set.

---

* Assistant Professor, School of Business Administration, Kwansei Gakuin University.
The author would like to express his deep appreciation to Professor Lee G. Cooper and Professor Dominique M. Hanssens at UCLA, and Professor Masao Nakanishi at Kwansei Gakuin University in Japan for their many helpful comments.

Understanding the competitive-market structure (CMS) is essential to the development of effective marketing strategy and plans, new-product development, promotion-mix design, product management, and so on. On the basis of CMS analyses, we try to identify which brands are competing more intensively or how subgroups of brands are competing with others in terms of marketing instruments and attributes.

The CMS-analysis methods, according to Day, Shocker, and Srivastava (1979), Bourgeois, Haines, and Sommers (1987), and Shocker, Stewart, and Zahorik (1990), vary in terms of the following four aspects: behavioral or judgmental data, supply- or demand-oriented measures, confirmatory or exploratory analysis, and spatial or non-spatial representation of CMS. In this paper, we present a system of *non-spatial* CMS models that attempt to *test* CMS hypotheses on the basis of *demand*-oriented *behavioral* data, *i.e.*, brand-switching matrices.

Brand-switching matrices are a major type of data used for CMS analyses (*e.g.*, Kalwani and Morrison 1977; Rao and Savabala 1981; Harshman, Green, Wind and Lundy 1982; Grover and Dillon 1985; Novak and Stangor 1987; Grover and Srinivasan 1987; Colombo and Morrison 1989; Kumar and Sashi 1989; Novak 1993; Cooper and Inoue 1996). Among the CMS analyses proposed for brand-switching data are the two-way latent-class analysis[1] (*cf.* Goodman 1974). The models of Grover and Dillon[2], Grover and Srinivasan, and Colombo and Morrison are the straightforward extension of latent-class models. We infer CMSs on the basis of parameter estimates, *i.e.*, latent class probabilities and conditional choice probabilities. However, because we cannot always identify latent-class models, that is, because parameters for the models are not always uniquely identified, we might derive a very different CMS. The Grover and Srinivasan (GS) approach is *not* the exception. Prior to the MIGHT system, no CMS models based on latent-class models present the identification *a priori and theoretically based on the population.*

The latent-class analysis attempts to decompose the switching-probability matrix $P=\{p_{ij}\}$ into K latent classes within which variables are assumed to be locally-independent, *i.e.*, the rows and columns are independent within latent classes. With the assumption of the zero-order stochastic process, we can formulate the latent-class model as $p_{ij} = \sum_{k=1}^{K} \pi_{ik}\pi_{jk}w_k$ and by using matrix-representation

$$P = \Pi W \Pi' \qquad (1)$$

where $\Pi = \{\pi_{ik}\}$ in which π_{ik} is the conditional probability that brand i is chosen at latent class k (i=1,...,B; k=1,...,K) and $W = diag\{w_k\}$ in which w_k is the probability that an individual will be in latent class k.

1 Hereafter the latent-class models indicate the two-way models without the term "two-way."

2 Grover and Dillon's model is based on, not a two-way, but a three-way switching matrix. However, it is still a straightforward application of the latent-class model.

A brief example illustrates the main point of this paper. Because the GS approach does not guarantee the identification *a priori* and theoretically, two substantially different CMSs can appear, which feature the value of the identification of the MIGHT system. Suppose we have a switching matrix composed of four brands over two purchase occasions. On the basis of the procedure suggested by GS, *i.e.*, the sequential estimation of parameters from K=1 to an arbitrary value of K while restricting "nearly-zero" values to zero with different starting values, we get the following set of parameter estimates:

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & 0 & .98 & .27 \\ 0 & 1 & 0 & 0 & 0 & .38 \\ 0 & 0 & 1 & 0 & .02 & .18 \\ 0 & 0 & 0 & 1 & 0 & .17 \end{pmatrix} \text{ and } W = diag\ \{0, 0, .06, .09, .20, .65\}.$$

The χ^2 statistic of the above model with seven parameters is 18.25, R^2 is .89, and $\bar{R}^2$ (adjusted R^2) is .84.[3] Regarding these statistics, see equations (4a) through (4g) presented in a later section. We infer the following CMS: Brands 1 and 2 do not have loyal-segments. The second switching submarket accounts for the largest proportion in the product class (.65) and all brands compete in this switching submarket. In comparison, in the first switching submarket only brands 1 and 3 compete each other with the dominance of brand 1.

In contrast, we get the following solution that shows the same level of goodness-of-fit with the same number of parameters as the above, that is, the χ^2 statistic is 17.88, R^2 is .89, and $\bar{R}^2$ is .84:

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & 0 & .14 & .34 \\ 0 & 1 & 0 & 0 & 0 & .46 \\ 0 & 0 & 1 & 0 & .12 & .20 \\ 0 & 0 & 0 & 1 & .74 & 0 \end{pmatrix} \text{ and } W = diag\ \{.18, 0, .07, 0, .23, .52\}.$$

On the basis of the above result, we infer the following CMS: Brand 2 and 4 do not have loyal-segments. Brands 1, 3, and 4 compete with the others in the first switching submarket with the dominance of brand 4. Brands 1, 2, and 3 compete with the other in the second, largest submarket (.52). There is no submarket in which all brands compete.

The above results are derived for a switching-frequency-matrix, N=$\{n_{ij}\}$ where n_{ij} is the number of consumers who bought brand i at time t and brand j at time t+1, based on a set of

---

3 For the four brand-loyal and one switching submarket solution, the χ^2 statistic is 19.19, R^2 is .88, and $\bar{R}^2$ is .80 and for the four brand-loyal and three switching submarket solution, the χ^2 statistic is 18.25, R^2 is .89, and $\bar{R}^2$ is .81.

the following population-parameter values, which *are different* from the above:

$$N=\begin{pmatrix}93 & 21 & 25 & 8\\ 25 & 38 & 25 & 12\\ 14 & 18 & 33 & 8\\ 9 & 16 & 12 & 43\end{pmatrix},\quad \Pi=\begin{pmatrix}1 & 0 & 0 & 0 & .5 & .1\\ 0 & 1 & 0 & 0 & .2 & .4\\ 0 & 0 & 1 & 0 & .2 & .3\\ 0 & 0 & 0 & 1 & .1 & .2\end{pmatrix}, \text{and } W=diag\{.10,.05,.05,.10,.40,.30\}$$

We should note two points from the above example. One is that the lack of identification of models can lead us into an incorrect implication of CMS which, in turn, provokes mis-specification of competitive marketing strategy and plans. Hence, we need to develop a model of which identification is ensured *a priori* and theoretically. Equally importantly, the second point is that the GS approach cannot be relied on to discover CMS.

In sum, the identification of latent-class models is critical from the viewpoint of CMS analysis since the inference from one solution might be substantially different from another solution and further might be incorrect or *not the true* structure. The failure of the identification misleads to competitive marketing strategy and plans. Thus, we need to impose some meaningful restrictions *a priori* on latent-class models to ensure their identification. The purpose of this paper is to present a system of latent-class models, the MIGHT system, that secures the identification *a priori and theoretically based on the population* and to show the superiority of the *a priori*, theoretically identified model to other models, especially the GS model, of which identification is not assured *a priori* and theoretically. The following are to be presented.

1. We introduce a system of non-spatial CMS models, which attempts to test CMS hypotheses based on brand-switching matrices.
2. The MIGHT system, composed of four submodels, ensures the identification *a priori and theoretically based on the population*, rather than on data.
3. We compare the MIGHT system with the GS model and show the superiority of the MIGHT system in terms of the goodness-of-fit, the ease of interpretation, and the stability of parameter estimates.
4. We also compare the MIGHT system with other CMS models and demonstrate its greater flexibility and lesser restrictiveness.

The remainder of the paper is organized as follows. In the next section, we present the MIGHT system and discuss its assumptions and identification. In the third section, we compare the MIGHT system with the GS model, which is a well-known CMS model based on latent-class models. Then, we compare conceptually the MIGHT system with other models. Finally, we conclude this paper by addressing future research.

2. The MIGHT System

The MIGHT system consists of four submodels that differ with respect to four assumptions which provide the identifiability of models *a priori and theoretically based on the population*. The four assumptions are the zero-order stochastic choice process, the existence of a universal submarket to which all brands belong, the exclusive/overlapped submarket, and the homogeneous/heterogeneous attractions. The first two assumptions hold for all models and the different combinations of the last two assumptions create four different models. The MIGHT system is a collection of CMS testing models based on latent-class models. We begin this section by discussing the issue of the identification. Then, we introduce four submodels of the MIGHT system and the steps for testing CMSs. Finally, we show the stability of the MIGHT system.

A Priori and Theoretical Identification Based on the Population

We cannot always identify latent-class models. The necessary and sufficient condition for the identification of a latent-class model is that the first-order derivative matrix, of which each element is a partial derivative of the joint probability with respect to conditional choice probability or latent probability, is a full column rank (Goodman 1974). In other words, the rank of the first-order derivative matrix is equal to the number of columns. In our context, each element of matrix is $\partial p_{ij} / \partial \pi_{ik}$ or $\partial p_{ij} / \partial w_k$.

A substantial amount of calculus is required to ascertain whether a specific latent-class model is identifiable *a priori* and theoretically. The identification depends on the choice of model and on the specification of fixed, constrained, and free parameters. Thus, the examination of the identification requires case-by-case calculus to see if the first-order derivative matrix is a full column rank. With regard to the model, if a user changes a constraint on a model, s/he needs to examine the column rank from the beginning. We have only one study that investigates the identifiability, *i.e.*, McHugh (1956). However, McHugh examines the identification only for a simplest latent-class model composed of only dichotomous variables. Generally, the scrutiny of identification requires a considerable amount of calculus labor.

Because of this difficulty, we, as well as GS, often employ a data-base approach that uses estimated values to assess identifiability (*e.g.*, MLLSA supplied by Clogg 1977). However, this approach is no more than *post hoc* and has no rationale behind it. Bentler (1980) strongly criticizes this approach as theoretically unsound since the identification is a problem of *population*, independent of sampling considerations. Hence, logically speaking, we should first assess the identification of a model *a priori* and theoretically, then estimate the parameters of the model. However, the data-base approach follows the reverse procedure. Consequently, it is valuable to develop CMS models based on latent-class models that establish their identifiability *a priori and theoretically based on the population.*

Assumptions of the MIGHT System

The MIGHT system assures the identifiability of a model *a priori and theoretically based on the population* by incorporating four assumptions with regard to the order of consumer stochastic choice behavior and heterogeneity in both choice behavior and consideration sets among submarkets. First, we examine the order of the stochastic choice process. We assume the zero-order stochastic choice process in the MIGHT system as well as GS do in their model. That is, a consumer's brand choice probabilities at a purchase occasion are unaffected by his/her previous purchase history. Bass, Givon, Kalwani, Reibstein, and Wright (1984) investigate the order of the stochastic choice process with respect to some categories of frequently purchased low-price packaged products. Their study indicates that the purchase sequences of a majority of stationary consumers are consistent with the zero-order assumption.

We next consider the heterogeneity of consumer choice behavior. In essence we have two approaches to deal with it. One approach is a conjugate distribution method, such as the beta-binomial model (*e.g.*, Sabavala and Morrison 1977) or the Dirichlet-multinomial model (*e.g.*, Jain, Bass, and Chen 1990). The other approach is the latent-class mixture model (*e.g.*, Kamakura and Russell 1989). In the former approach the conjugate distributions account for the heterogeneity and in the latter a couple of mass-points do. For the MIGHT system we apply the latter approach and presume the latent-class mixture attraction model (*e.g.*, Cooper and Nakanishi 1988) on the basis of the random-utility model. We model two types of attraction: homogeneous and heterogeneous attraction across submarkets. Models of the homogeneous attractions turn out to be the same as the restriction of the proportional conditional probabilities across submarkets in the context of latent-class models. In other words, we restrict the ratio of π_{ik} / π_{jk} to be equal to π_{ih} / π_{jh} $(i,j \in h,k;\ h \neq k)$ under the assumption of homogeneity.

Finally, we think about the heterogeneity of consideration sets among submarkets. There are three general ways that consideration sets may differ over submarkets. First, each submarket can have its exclusive set of brands (*e.g.*, caffeinated vs decaffeinated coffees). Second, there can be overlap of brands in the submarkets. Third, there can be submarkets in which all consumers consider all brands (universal submarkets). In the MIGHT system, we consider two combinations of the above, that is, either the existence of one universal and some exclusive submarkets or that of one universal and some overlapped submarkets. The specification with exclusive submarkets turns out to be equivalent to restricting the choice probability vector with respect to submarket S_k, Π_k ($1 \le k \le K$), to be pairwise orthogonal or $\Pi_h'\Pi_k=0$ ($\forall h \neq k$; $h,k \geq 1$) in latent-class models.

We admit the existence of a universal submarket for the following substantive, theoretical, and empirical reasons even though the number of brands in consideration sets, according to Hauser and Wernerfelt (1990), varies from two to eight.

1. *A universal submarket assimilates some types of uncertainty involved in consideration sets.* First, the consideration sets, even for a customer, can be changing over time depending on purchase occasions (*e.g.*, Siddarth, Bucklin, and Morrison 1993). A universal submarket can assimilate such dynamism. Second, we should recognize the distinction among evoked sets, consideration sets, choice sets, and purchased sets (*e.g.*, Shocker, Ben-Akiva, Boccara, and Nedungadi 1991). We require a survey to know about the first three concepts and purchase-history data tell us only about the last concept. The model is based on the consideration sets, not on the purchased sets. A universal submarket can grasp the ambiguity in the consideration sets.

2. *The existence of a universal submarket provides us with a basis for the null CMS hypothesis that a market is composed of only one universal submarket.* This null hypothesis is the same as one used by Urban, Johnson, and Hauser (1984), *i.e.*, the aggregate constant ratio (ACR) model.

3. *We should statistically investigate its existence.* In other words, if we do not have such a universal submarket, the latent-class probability should not be significantly different from zero and, otherwise, it should be significant.

Regarding the last two assumptions on the heterogeneity in the choice behavior and consideration sets among submarkets, the assumption of exclusive submarkets is more restrictive, but more parsimonious than that of overlapped submarkets, as is the assumption of homogeneous attractions than that of heterogeneous attractions. However, whether these restrictive assumptions are valid is a matter of statistical tests. That is, if truly exclusive submarkets exist in a particular application, the tests would not reject the assumption of the exclusiveness. Similarly, if there exist homogeneous attractions in a particular application, the tests would not reject the assumption of the homogeneity. Hence, MIGHT 1 might be accepted rather than MIGHT 2 or 3 and MIGHT 2 might be accepted rather than 3, depending on the application. As we see in the application section, based upon information criteria, we chose MIGHT 3 instead of MIGHT 1 or MIGHT 2.

In summary, the MIGHT system makes four assumptions so as to assure *a priori*, theoretical identifiability: the zero-order stochastic choice process, the existence of a universal submarket, the heterogeneous/homogeneous attractions, and the exclusive/overlapped consideration sets. The first and second assumptions hold for all models of the

MIGHT system, so that we have four phases of the MIGHT system, each of which has a different combination of two levels of the third and fourth assumptions. To begin, we explain the full model, consisting of the least restrictive assumptions.

Full Model (MIGHT 4)

The full model is based on the following assumptions.

(A 1) The brands offered in a market are divided into K submarkets, expressed as S_k (k=1, 2, ..., K).[4] We denote the relative size of a submarket S_k as w_k.

(A 2) There is an additional submarket S_0, universal submarket, in which all brands offered in the market are included. Its relative size is denoted by w_0. Thus, $w_0 = 1 - \sum_{k=1}^{K} w_k$.

(A 3) In submarket S_k (k=0, 1, ..., K), the deterministic attraction of brand i in submarket k is α_{ik} and the corresponding random component is ε_{ik}, independently and identically distributed as the type-II extreme-value density. A consumer chooses a brand with the maximum of random attraction, $\alpha_{ik} + \varepsilon_{ik}$. The probability that a consumer chooses product i ($i \in S_k$) is specified as

$$\pi_{ik} = \alpha_{ik} \,/\, \sum_{j \in S_k} \alpha_{jk} \,, \tag{2}$$

otherwise 0.

(A 4) The choice probability follows the zero-order process.

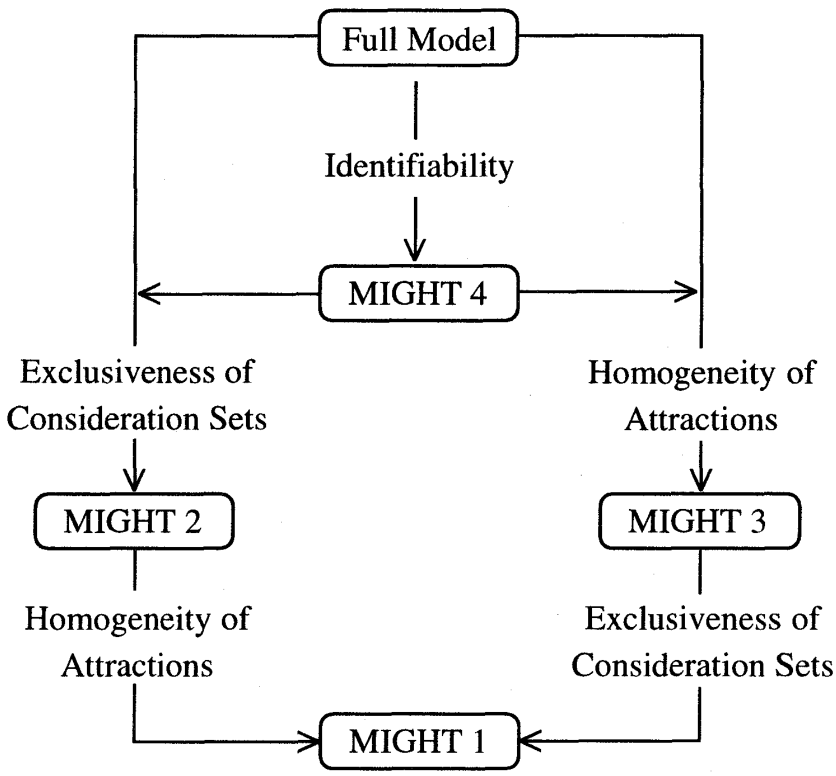
We define an unstructured market in the same way as Urban, Johnson, and Hauser do in the PRODEGY model. That is, the unstructured market consists only of the submarket S_0 with the restriction w_0=1. If a market has any other competitive market structure, then, in addition to the submarket S_0, the submarkets S_k ($k \geq 1$) exist in the market, or w_0<1. Thus, we can test CMS hypotheses against the null hypothesis with the restriction w_0=1.

The full model does not restrict the attractions or the consideration sets. The zero-order assumption *per se*, however, is not enough to secure *a priori*, theoretical identifiability of the full model. Hence, we have to examine the identifiability of MIGHT 4 models individually, which requires as much case-by-case calculus as general latent-class models. Consequently we need additional assumptions to assure *a priori*, theoretical identifiability of models. We

4 We use S_k to denote both a submarket and a segment of buyers, in the sense that each buyer is associated with the subset of products/brands, consideration set, which s/he most prefers. It should be noted that the MIGHT system is able to assign consumers into submarkets *posteriori*, as well as Grover and Srinivasan's (1987) model. In this sense, the MIGHT system is capable of the simultaneous analysis of competitive market structuring and market segmentation and so is Grover and Srinivasan's model.

impose two further restrictions on the full model, concerning the heterogeneity of consumer choice behavior, *i.e.*, the exclusiveness of submarkets and/or the homogeneity of attractions among submarkets. MIGHT 3, 2, and 1 differ from each other in the combination of the assumptions dealing with the submarkets (exclusive/overlapped) and the attractions (homogeneous/heterogeneous), on which choice probability is based. Figure 1 depicts the relationship among models of the MIGHT system.

Figure 1 Relationship among the Models in the MIGHT System



MIGHT 4 is the full model (general latent-class model) where we ensure the identification. MIGHT 3 is a special case of MIGHT 4 in which we restrict attractions to be homogeneous. MIGHT 2 is a special case of MIGHT 4 where we impose the exclusiveness on submarkets. MIGHT 1 is a special case of MIGHT 3 with exclusive submarkets or, alternatively, that of MIGHT 2 with homogeneous attractions. Because of the complex calculus involved to see the identification of MIGHT 4, we limit our discussion to MIGHT 1, 2, and 3.

MIGHT 3

MIGHT 3 is based on the following three assumptions:

(A 1) The same assumption as that of the full model.

(A 2) The same assumption as that of the full model.

(A 3) In any submarket S_k (k=0, 1, ..., K), the deterministic attraction of brand i in submarket k is α_i and the corresponding random component is ε_i, independently and identically distributed as the type-II extreme-value density. A consumer chooses a brand with the maximum of random attraction, $\alpha_i + \varepsilon_i$. The probability that a consumer chooses product i ($i \in S_k$) is specified as

$$\pi_{ik} = \alpha_i / \sum_{j \in S_k} \alpha_j , \qquad (3),$$

otherwise 0.

(A 4) The same assumption as that of the full model.

MIGHT 3 allows submarkets to have overlapping membership, but restricts the attractions to be homogeneous across submarkets. With regard to MIGHT 3, the choice probability to be estimated is only $\{\pi_{i0}\}$. A brand belonging only to the submarket S_0 and a submarket S_k is referred to as *SkS* (Submarket k Specific) brand. Then, given that the number of parameters to be estimated is smaller than $[B\times(B+1)/2]-1$ where B is the number of brands, *MIGHT 3 is identified whenever the number of SkS brands for a submarket is greater than or equal to two* (see appendix). We can satisfy this condition quite easily in most empirical applications, so that MIGHT 3 achieves identifiability for nearly all cases.

MIGHT 2

MIGHT 2 is based on the following assumptions:

(A 1) The brands offered in a market are divided into K submarkets, expressed as S_k (k=0, 1, ..., K). However, submarkets are mutually exclusive. In other words, customers who belong to a submarket S_k never choose any of the brands assigned to another submarket S_h ($h \neq k$). We denote the relative size of a submarket S_k as w_k.

(A 2) The same assumption as that of the full model.

(A 3) The same assumption as that of the full model.

(A 4) The same assumption as that of the full model.

As a consequence, MIGHT 2 imposes exclusiveness of submarkets on the full model. Interestingly, *MIGHT 2 is always identifiable if the number of brands is greater than two* (see appendix).

MIGHT 1

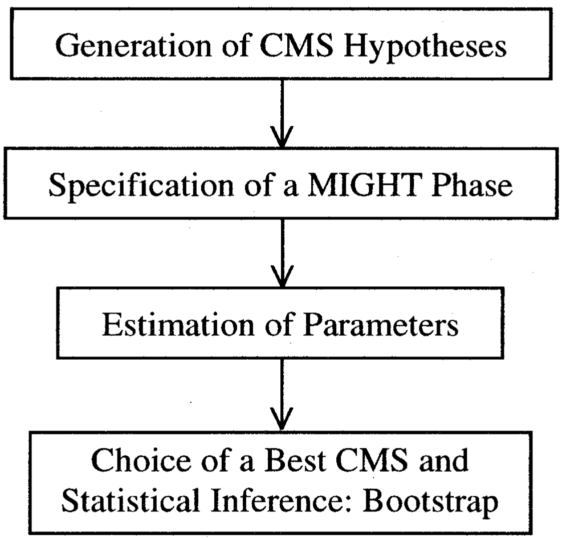
MIGHT 1 assumes the following:

(A 1) The same assumption as that of MIGHT 2.
(A 2) The same assumption as that of the full model.
(A 3) The same assumption as that of MIGHT 3.
(A 4) The same assumption as that of the full model.

Inoue and Nakanishi (1991) show $w_k = (1 - w_0) \sum_{j \in S_k} m_i$ from the restriction $\sum_{all\ j} p_{ij} = m_i = \alpha_i / \sum_{all\ j} \alpha_j$. Consequently the parameters to be estimated are only $\{\pi_{i0}\}$ and w_0. *MIGHT 1 is identifiable for any case* (see appendix).

Testing Competitive-Market-Structure Hypotheses and Choice of a Best Structure

Figure 2 Steps of Testing CMS Hypotheses



The procedure of testing CMS consists of four steps (Figure 2). The first step is to generate a set of CMS hypotheses. The second step is to specify a certain phase of the MIGHT system, *i.e.*, either MIGHT 1, 2, or 3. The third step is to estimate parameters under those CMSs and calculate statistical criteria. The fourth step is to perform the bootstrap to choose a best CMS hypothesis among the null and alternative CMS hypotheses and to infer the statistical property of parameter estimates. At the present, the first two steps require various aspects, such as Managers' experience or prior knowledge. We can count on other CMS models. Alternatively, we can sequentially apply the varimax rotation to residual matrices for an exploratory purpose as discussed in Inoue (1993).

Once we specify a particular CMS and phase of the MIGHT system, we estimate parameters by using the EM algorithm (Goodman 1974; Dempster, Laird, and Rubin 1977; Mooijaart and van der Heijden 1992). We estimate the parameters of the MIGHT models with restrictions corresponding to their assumptions. We impose the proportionality on conditional probabilities regarding MIGHT 1 and 3, *i.e.*, $\pi_{ik} / \pi_{jk} = \pi_{ih} / \pi_{jh}$ $(h \neq k)$. We impose the orthogonality on conditional probabilities regarding MIGHT 1 and 2, *i.e.*, $\Pi_h' \Pi_k = 0$ $(\forall h \neq k;\ h,k \geq 1)$.

The usual likelihood-based method to infer statistical properties of parameters, such as an information matrix or likelihood-ratio test, does not work for the latent-class mixture models because of the violation of the regularity conditions (Titterington, Smith, and Makov 1985; McLachlan and Basford 1988). Thus, special cases of latent-class models, such as Grover and Dillon or Grover and Srinivasan, and the latent-class mixture models such as the latent class BTL (binary logit) model (Dillon, Kumar, and de Borrero 1993), the latent class MNL model (Zenor and Srivastava 1993), the latent class censored regression (Jedidi, Ramaswamy, and DeSarbo 1993), and the latent class ML INDSCAL (CLASCAL) model (Winsberg and De Soete 1993) cannot rely on likelihood-based tests.

We can, however, use the bootstrap (Efron 1979; Efron and Tibshirani 1993). On the basis of the bootstrap, we can choose a best CMS, together with a particular phase of the MIGHT system, among the null and alternative CMS hypotheses and infer the statistical properties of parameters. However, judging from the approximate expression discussed in Titterington, Smith, and Makov, and McLachlan and Basford, the correction factor approaches 1 as the total number of observations gets large. Hence, we can conjecture that, in large samples, we can use regular likelihood-based methods for inferring parameters and choosing a best CMS hypothesis on the basis of the following criteria (*e.g.*, Bozdogan 1987):

$$LR = 2 \sum_{i=1}^{B} \sum_{j=1}^{B} n_{ij} \log \left(\frac{n_{ij}}{\hat{n}_{ij}} \right). \tag{4a}$$

$$\chi^2 = 2 \sum_{i=1}^{B} \sum_{j=1}^{B} \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}. \tag{4b}$$

$$AIC = LR + 2NP. \tag{4c}$$

$$CAIC = LR + NP(\ln n + 1) \tag{4d}$$

$$SBIC = LR + NP \times \ln n\ . \tag{4e}$$

Also, we can formulate two normalized measures bounded between 0-1 in which LR in the former is not adjusted, but LR in the latter is adjusted by its degrees-of-freedom ($\overline{LR}=LR/df$) as follows:

$$R^2 = \frac{LR_{null} - LR}{LR_{null}} \tag{4f}$$

$$\bar{R}^2 = \frac{\overline{LR}_{null} - \overline{LR}}{\overline{LR}_{null}} \tag{4g}$$

Stability of the MIGHT System: Identifiability, Estimation Accuracy, Robustness, and Distinguishability

Inoue (1993) investigated the stability of the MIGHT system based on two simulation experiments. The first study shows the satisfactory accuracy of parameter estimates, the fairly robustness, and the adequate identifiability of all MIGHT models. Importantly we find that MIGHT 3 is superior to MIGHT 1 and 2 in all the three aspects. The second simulation compares *between* MIGHT models. Comparative simulations are needed since the MIGHT models are applied as a system, rather than as a discrete collection of four separate models. We discover that MIGHT models are distinguishable with a substantial level of the estimation accuracy, robustness, and the identifiability as shown in the first simulation.

In sum, from the findings of the two simulation studies, we have verified the adequate level of the stability of the MIGHT system with respect to the identification of a true CMS, the estimation of parameters, the robustness, and the distinguishability of MIGHT models.

3. Comparison with Grover and Srinivasan's Approach

Grover and Srinivasan's Approach

Grover and Srinivasan (1987) introduced a simultaneous approach to market segmentation and market structuring based on latent-class models. They assume a zero-order choice-process and decompose a market into two segments, *i.e.*, B brand-loyal (BL) segments and M brand-switching (BS) segments where B is the total number of brands in the market. Consumers in BL segments are supposed to purchase only one brand per segment, in comparison to consumers in BS segments who take into consideration subsets of brands specific to each BS segment. Thus, we get

$$P_{ij} = \sum_{k=1}^{B} V_k + \sum_{k=B+1}^{B+M} \pi_{ik}\pi_{jk} w_k \tag{5}$$

Using a matrix-notation in equation (1), we can specify Π as $\{I_{BL} \mid \pi_{BS}\}$ in which I_{BL} is the identity matrix with B dimensions, $\pi_{BS} = \{\pi_1 \mid \dots \pi_k \dots \mid \pi_M\}$, where $\pi_k = \{\pi_{ik}\}$ is a column vector of conditional choice probabilities given submarket k, where $\pi_{ik} = 0$ if brand i does not belong to submarket k (BS segment), and $\boldsymbol{W}$ as diag$\{V_k \mid w_k\}$ in which V_k and w_k are the sizes of BL and BS segment k.

The procedure to determine market segments and CMSs suggested by GS is as follows: calculate the $\bar{R}^2$'s by varying the number of BS segments from one to a certain number and decide the number of BS segments at the point of substantial improvement in terms of the goodness-of-fit measures. In that calculation process, they estimate the parameters of the model with B plus M submarkets, while restricting π_{ik}'s equal to zero, of which estimates are less than .01 for the purpose of identification and re-estimating the model under the restriction. Several starting values are used to maximize the chances of the EM solutions ending with global-optimum parameter estimates. Again, the restrictions are just *post hoc* and do *not* have any logic for the assurance of the identification theoretically based on the population.

Application of MIGHT System to Grover and Srinivasan's Data

We apply the MIGHT system to the data given in Table 1. First we need to construct a set of CMS hypotheses and specify a certain phase of the MIGHT models for each of the hypotheses. Table 2 shows CMS hypotheses tested to the data in Table 1. The second row indicates the CMSs and the third row the phases of the MIGHT system. The difference between hypotheses 1 and 2 is that the former is estimated under "brand-primary CMS" with MIGHT 1 but the latter with MIGHT 2. The numbers in the fourth row indicate submarkets to which brands pertain, where we remove 0s. However, it should be kept in mind that, from the assumption, all brands belong to submarket S_0. As an instance, under hypothesis 5, brand 1 belongs to its BL submarket 1 and to BS submarket 13, but brand 7 only to its BL submarket 7.

We construct hypothesis 5 based on the "underlined" estimates of GS's solution (the first and second largest conditional-probabilities). We create hypothesis 6 based on their statistically significant "bold" estimates (conditional probabilities of which ratio of estimate to standard error is greater than 2). Hypotheses 7 and 8 are similar to hypotheses 5 and 6 (GS's outcome), but we accentuate the CMSs in such a way that they are easier to interpret.

Table 1 Instant Coffee Cross-Classification Matrix

	Br1	Br2	Br3	Br4	Br5	Br6	Br7	Br8	Br9	Br10	Br11
Br1	93	7	17	19	18	43	1	4	6	7	10
Br2	9	80	12	11	24	7	4	2	6	3	3
Br3	9	14	46	3	7	7	4	2	2	0	9
Br4	19	18	4	82	29	14	0	4	9	2	6
Br5	26	11	6	35	184	24	3	11	18	6	6
Br6	15	13	8	13	28	127	4	3	3	8	8
Br7	2	0	3	2	1	7	17	3	0	1	4
Br8	4	3	4	3	6	5	2	27	1	0	4
Br9	5	3	2	4	16	4	0	1	46	9	2
Br10	6	1	4	1	5	9	0	0	11	15	2
Br11	10	4	4	4	2	10	2	2	5	2	27

* Citation from Grover and Srinivasan (1987)

Br1	: High Point	Decaffeinated	Regular (spray dried)	Procter & Gamble
Br2	: Taster's Choice	Caffeinated	Freeze dried	Nestlé
Br3	: Taster's Choice	Decaffeinated	Freeze dried	Nestlé
Br4	: Folgers	Caffeinated	Regular (spray dried)	Procter & Gamble
Br5	: Maxwell House	Caffeinated	Regular (spray dried)	General Foods
Br6	: Sanka	Decaffeinated	Regular (spray dried)	General Foods
Br7	: Sanka	Decaffeinated	Freeze dried	General Foods
Br8	: Maxwell House	Caffeinated	Freeze dried	General Foods
Br9	: Nescafé	Caffeinated	Regular (spray dried)	Nestlé
Br10	: Nescafé	Decaffeinated	Regular (spray dried)	Nestlé
Br11	: Brim	Decaffeinated	Freeze dried	General Foods

Table 2 Tested CMS Hypotheses and Result of Application of MIGHT System to Data in Grover and Srinivasan

	Hypothesis1	Hypothesis2	Hypothesis3	Hypothesis4	Hypothesis5	Hypothesis6	Hypothesis7	Hypothesis8
CMS	Brand	Brand	Caf/Decaf	Regular/FD	GS1	GS2	Compound2	Compound1
Model	1	2	2	2	3	3	3	3
Br 1	1	1	2	1	1•13	1•12•13•14	1•18	1
Br 2	2	2	1	2	2•14	2•12•14	2•12•13•16	2•12•13•16
Br 3	2	2	2	2	3•14	3•14•15	3•13•16	3•13•16
Br 4	3	3	1	1	4•12	4•12•13•14	4•12•18	4•12
Br 5	4	4	1	1	5•12	5•12•13•14	5•12•15•18	5•12•15
Br 6	5	5	2	1	6•13	6•13•14•15	6•18	6
Br 7	5	5	2	2	7	7•13•14	7•13	7•13
Br 8	4	4	1	2	8	8•12•13•14	8•12•13	8•12•13
Br 9	6	6	1	1	9•15	9•12•15	9•12•14•15•17	9•12•14•15
Br 10	6	6	2	1	10•15	10•13•15	10•14•15	14•15
Br 11	7	7	2	2	11	11•13•14	11•13•17	11•13
AIC	482.5	453.9	1004.3	945.2	226.2	296.2	174.2	171.7
CAIC	552.3	587.2	1137.6	1078.5	384.9	454.9	351.9	330.4
SBIC	541.3	566.2	1116.6	1057.5	359.9	429.9	323.9	305.4
R2	.70	.73	.37	.41	.88	.84	.92	.92

That is, submarket 12 contains all caffeinated brands, submarket 13 involves all freeze dried brands, submarket 14 includes all Nescafé brands, submarket 15 contains only Nescafé and Maxwell House regular coffee, and submarket 16 involves only Taster's Choice brands. The distinction between hypotheses 7 and 8 is that the former has a BL segment for brand 10 and two additional submarkets, but the latter does not. Submarket 17 contains Nescafé caffeinated and Brim coffee and submarket 18 contains four regular coffees marketed by two major manufacturers (Procter & Gamble and General Foods). The fifth row shows the criteria to choose a best CMS, such as AIC, CAIC, SBIC, and R^2. It is important to note that hypothesis 8 is identifiable.

From GS's results, we already know that there exists a CMS. This is also confirmed from a large *LR* of the null hypothesis (1530.0). We conclude that hypothesis 8 is the best CMS on the basis of all criteria. It should be noted that the χ^2 under hypothesis 8 was 117.8 (95 degrees of freedom and p=.06) with a much fewer number of parameters than GS's results (*LR*=100.8 with 77 degrees of freedom; p=.04). Thus, the MIGHT system does not reject the CMS hypothesis at the 5% level, but GS's results rejected their CMS hypothesis at the 4% level. Further, AIC, CAIC, SBIC, R^2, and $\bar{R}^2$ for the GS model (MIGHT model) are 186.8 (171.7), 459.8 (330.4), 416.8 (305.4), .93 (.92), and .906 (.908) respectively. Hence, the MIGHT system achieves better results than the GS model in terms of all criteria except for the trivial difference in R^2.

Next, we examine the estimation result of the MIGHT system under hypothesis 8. Table 3 shows the estimated parameters and their t-values (in parentheses) under hypothesis 8. On the ground that the conditional probabilities reflect the rational relationships of attractions, we conclude that brand 1 has the largest attraction and brand 7 the least attraction in the market. Submarket-size parameters are presented in the second row. The first is the size of the universal submarket and the next 10 are estimates of sizes of 10 brand-loyal submarkets corresponding to V_k in the GS model. Readers should note that brand 10 does not have its own loyal-segment that is statistically significant. The remaining five parameters are the sizes of switching segments.

We thus ascertain that, in addition to the largest submarket S_0 (w_0=.45) and 10 BL submarkets (no BL submarket for brand 10), there are 5 BS submarkets, instead of 4 (GS's outcome). The first BS submarket S_{12} comprises all caffeinated brands supplied by two major brands, *i.e.*, Procter & Gamble and General Foods (similar to submarket S_{12} in the GS model). This submarket is the second largest (w_{12}=.12), so that for those two marketers, this submarket is critical. As stated in note 3, the MIGHT system is able to detect a group of consumers who compose a particular submarket. Hence, this tracking approach for this submarket might be useful especially for two major or five caffeinated brands. In comparison, all freeze-dried coffees belong to submarket S_{13}. Submarket S_{14} consists of only

Table 3 Result of Estimation of the MIGHT System under Hypothesis 8

Conditional Probabilities										
π_{10}	π_{20}	π_{30}	π_{40}	π_{50}	π_{60}	π_{70}	π_{80}	π_{90}	$\pi_{10,0}$	$\pi_{11,0}$
.26 (5.90)	.06 (4.24)	.06 (4.58)	.11 (4.99)	.16 (6.70)	.18 (8.64)	.02 (3.53)	.03 (4.03)	.03 (4.82)	.04 (3.82)	.06 (6.46)

Universal Submarket	*Brand-loyal Submarkets*										*Switching Submarkets*				
w_0	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{11}	w_{12}	w_{13}	w_{14}	w_{15}	w_{16}
.45 (17.7)	.03 (3.00)	.04 (5.60)	.02 (4.13)	.04 (5.64)	.08 (6.08)	.07 (8.54)	.01 (3.93)	.02 (4.43)	.02 (4.12)	.01 (4.80)	.12 (4.81)	.03 (3.36)	.02 (4.05)	.02 (1.35)	.02 (2.58)

two Nescafé brands, which compete with Maxwell House also in submarket S_{15} (similar to submarket S_{15} in the GS model). Two Taster's Choice brands compete with each other in submarket S_{16} (similar to submarket S_{14} in the GS model). Accordingly, we observe a cannibalization in submarkets S_{14} and S_{16}. With regard to Nescafé and Taster's Choice, we suggest that these companies reposition their brands so as to avoid this phenomenon.

Comparison and Summary

Because GS did not present the data used for calibration in their paper, we compare the statistical properties of parameter estimates between the MIGHT system and the GS model. This comparison is done on the basis of the bootstrap, where we implement 50-times re-sampling from the data in Table 1. The parameters and their t-values (in parentheses) of the MIGHT system are given in Table 3 and those of the GS model in Table 4. We note that, since the restriction with respect to parameters for calibration was the same as that for model development, we can infer the statistical properties of parameters from Table 4.

It is remarkable that the estimates of the GS approach derived from the data in Table 1 are very different from those derived from the data for calibration (Table 3, page 147). More importantly, regarding the MIGHT system, 24 of 25 parameters are significant. In contrast, with respect to the GS model, only 13 of 39 parameters are significant, which the discrepancy of the estimates from the data in Table 1 with those from the calibration data, published in Table 3 on page 147, implies. This shows the superiority of the MIGHT system to the GS approach in terms of its stability, which underscores the risk caused by the lack of identification of the latter.

In comparison to the results of the GS model, the following four points are made:

1. The MIGHT system achieved a better result than the GS model in terms of AIC, CAIC,

Table 4 Result of Estimation of Grover and Srinivasan's Approach

	w_1	w_2	w_3	w_4
	.18	.05	.26	.17
	(2.31)	(.77)	(2.82)	(2.20)
V_k	π_{i1}	π_{i2}	π_{i3}	π_{i4}
.04	.01	.18	.25	.20
(2.57)	(.10)	(0.60)	(2.54)	(2.30)
.05	.11	0	.11	.07
(7.02)	(2.12)	(0.00)	(1.78)	(1.88)
.02	0	0	.15	.02
(2.91)	(0.00)	(0.00)	(1.29)	(0.68)
0	.36	.74	.06	0
(0.00)	(1.80)	(2.99)	(0.69)	(0.00)
.09	.33	.01	.13	.19
(3.98)	(2.95)	(0.01)	(2.18)	(2.93)
.07	.12	0	.13	.13
(9.53)	(2.00)	(0.00)	(2.59)	(3.30)
.01	0	.03	.05	0
(2.81)	(0.00)	(0.14)	(1.39)	(0.00)
.02	.03	0	.07	0
(3.59)	(1.36)	(0.00)	(1.26)	(0.00)
.03	.05	0	0	.14
(4.79)	(0.94)	(0.00)	(0.00)	(1.89)
.01	0	.01	0	.17
(1.31)	(0.00)	(0.12)	(0.00)	(2.23)
.02	0	.04	.07	.07
(5.39)	(0.00)	(0.55)	(2.10)	(2.26)

SBIC, and $\bar{R}^2$ under a CMS hypothesis for which identification was secured *a priori and* theoretically based on the population.

2. The significance level of the χ^2 statistic under the identifiable CMS hypothesis 8 with MIGHT 3 was greater than that of the GS model (.06 vs .04). In particular, the MIGHT system did not reject the null hypothesis.
3. The CMS implied by the MIGHT system accentuated the characteristics of competitive groups more than that of GS, so that the former was easier to interpret than the latter.
4. The parameters of the MIGHT system were more stable than those of the GS approach, which implies the risk caused by the lack of the identification of the latter.

Conceptual Comparison with Other Models

In this subsection, we compare conceptually the MIGHT system with four past studies in terms of three aspects: testing or detecting CMS, inclusion or exclusion of repeat purchases (diagonal elements of switching matrices), and exclusive or overlapped submarkets. The four models are Rao and Sabavala (1981), Urban, Johnson, and Hauser (1984), Grover and Dillon (1985), and Colombo and Morrison (1989). We denote them by RS, UJH, GD, and CM respectively. The RS model is, based on switching matrices, a CMS detection method where diagonal elements are fundamentally excluded from the analysis and submarkets are assumed to be exclusive. MIGHT 1 or MIGHT 2 is, in spirit, similar to the RS model even though the former can test CMS hypotheses and include the diagonal in analysis.

The UJH model is, based on forced-switching matrices[5], a CMS testing model where diagonal elements are primarily excluded and submarkets are assumed to be exclusive. MIGHT 1 is very close to the UJH model (Inoue and Nakanishi 1991). The null UJH model is MIGHT 1 with w_0=1 restriction. There are two important points. The first point is that we can extend the UJH model directly by applying the framework used in this paper (exclusive or overlapped submarkets and homogeneous or heterogeneous attractions). Consequently, we get the heterogeneous UJH model. The second point is that even MIGHT 1 solves most problems involved in the UJH model. That is, MIGHT 1 can deal with a singleton submarket that consists of only one brand. We can identify a best CMS, based on information criteria, among multiple hypotheses which are significantly different from the unstructured hypothesis.

The GD model is also, based on switching matrices, a CMS testing model where repeat purchases are not considered and submarkets are assumed to be exclusive. MIGHT 1 or MIGHT 2 is, in spirit, similar to the GD model even though the latter analyzes a three-way table of which the first two ways are brands and the last is switching occasions. However, ignoring repeat purchases seems to distort insights into CMS.

The CM model is, based on switching matrices, a CMS testing model where only two types of submarkets are taken into account: hard-core loyal and potential switchers. MIGHT 3, in which we assume that the size of loyal-submarkets is common among brands and no submarket but S_0 is assumed, is similar to the CM model. Thus, the CM is a more restrictive special case of MIGHT 1.

In sum, as we have seen, the MIGHT system captures the best aspects of the four past studies, but appears to be more flexible and less restrictive.

4. Conclusions and Future Research

In this paper, we introduced a system of non-spatial models of competitive-market-

5 Here forced switching matrices mean a type of switching matrices which are developed by methods such as experimentally-controlled forced switching incidence, preference ranks, or logit models.

structure analysis, the MIGHT system, which attempts to test CMS hypotheses based on latent-class models for the analysis of brand-switching matrices. From the viewpoint of designing marketing strategy and plans, the identification of models of competitive-market-structure analyses is critical since the lack of identification leads to incorrect design of them. The identification is an issue of population, not that of sampling, so that the identification should be ensured *a priori and theoretically based on the population.* However, prior to the MIGHT system no CMS model based on latent-class models presents the identification on the basis of the population. The MIGHT system ensures *a priori*, theoretical identification. The MIGHT system consists of four submodels that are different with respect to the order of consumer stochastic choice behavior and heterogeneity in both choice behavior and consideration sets among submarkets. The MIGHT system is compared with one well-known CMS model, Grover and Srinivasan's model, in detail and with other CMS models conceptually. The first comparison shows the superiority of the MIGHT system in terms of the goodness-of-fit, the ease of interpretation, and the stability of parameter estimates. The second comparison indicates its greater flexibility and lesser restrictiveness.

Future research in this area can take three directions. One is application studies of the MIGHT system. Inoue (1993) applied the MIGHT system to the Japanese beer market and derived some managerial implications. We need to examine the external validity and applicability of the MIGHT system based on a large number of the applications.

The second is asymmetrization so as to take into account the non-zero-order stochastic process. The symmetry in a theoretical switching matrix comes from the assumption of the zero-order process. Even though Bass *et al.* show empirical support for the assumption, it might still be restrictive. Cooper and Inoue (1996) deal with the asymmetry by employing a quasi-first-order specification.

The third direction is the incorporation of marketing-mix and/or product attributes. This can be done by expressing brand i's attractiveness as a function of marketing instruments and/or product attributes as we see in market-share modeling (*e.g.*, Cooper and Nakanishi 1988).

In sum, the MIGHT system, composed of four submodels, ensures identification *a priori and theoretically based on the population.* We hope that, even though *a priori*, theoretical identification of models is cumbersome to do, this important issue will continue to be investigated.

Appendix: The Identifiability of Each of The MIGHT Models

The necessary and sufficient condition for the identifiability of latent-class models, according to Goodman (1974), is that, for the first-order derivative matrix **D**, each element is

$\partial p_{ij} / \partial \pi_{ik}$ or $\partial p_{ij} / \partial w_k$, is a full column rank, *i.e.*, the rank of the first-order derivative matrix is equal to the number of columns. The MIGHT system assumes the zero-order process of choice probability. Consequently, the number of free p_{ij}'s is $[B\times(B+1)/2]-1$ where B is the number of brands, due to the symmetry of the joint probability matrix. Thus, the necessary condition is that the number of parameters is not greater than $[B\times(B+1)/2]-1$. Each brand is sorted by the submarket to which it belongs and it is supposed that the last brand B is the *SkS* brand of the submarket S_K. We denote the number of brands which belong to a submarket S_k by B_k, the number of *SkS* brands of the submarket by B_k*, and the group of the SkS brands by S_k*. We explicate the identifiability of MIGHT 3, 2, and 1 in succession.

Identifiability of MIGHT 3

The number of columns of matrix **D** for MIGHT 3 is $B+K-1$. We indicate the column vector of the first-order partial derivative as $\nabla^{i0}=\{\partial p_{ij} / \partial \pi_{i0}\}$ and $\nabla^{w_k}=\{\partial p_{ij} / \partial w_k\}$ of which the dimension is $[B\times(B+1)/2]-1$. The necessary and sufficient condition for the identifiability is that each vector is linearly independent or that there exists no non-trivial solution to the following equation:

$$\delta_{10}\nabla^{10} + \cdots + \delta_{B-1,0}\nabla^{B-1,0} + \delta_{w_0}\nabla^{w_0} + \cdots + \delta_{w_{K-1}}\nabla^{w_{K-1}} \qquad \text{(a1)}$$

We think about the two cases, i.e., $B_{K^*}\geq 2$ and $B_{K^*}=1$, separately. First, suppose $B_{K^*}=2$. Then, with respect to each p_{iB} $(1\leq i\leq B-1)$ row of matrix **D**, we can get the equation system of which the number of rows is $B-1$ and of which the number of parameters, δ_i's (i=10, ..., B–1 0, w_0), is B. Thus, it yields a ratio of δ_i to δ_j. We can choose i and j arbitrarily, so let $i=(B-1$ $0)$ and $j=w_0$. Keeping that in mind, we obtain $\delta_{B-1,0}=\delta_{w_0}=0$ from the $p_{B-1,B-1}$ $(B-1\in S_K)$ row. Then, focusing on each $p_{i,B-1}$, it gives $\delta_{i0}=0$ for all $i\in S_K$ $(0<k<K)$. Finally, it turns out that for all the parameters

$$\delta_{10} = \cdots = \delta_{B-1,0} = \delta_{w_0} = \delta_{w_{K-1}} = 0 \qquad \text{(a2)}$$

or we prove that each vector is linearly independent. We can prove this for the case $B_{K^*}>2$ in the same way. As a conclusion, MIGHT 3 is identifiable as long as it has a submarket composed of greater than or equal to two *SkS* brands.

Next, we consider the case where $B_{K^*}=1$. Moreover, assume $B_{k^*}=2$ for a submarket S_k $(0<k<K)$. As noted before, we acquire the equation system, of which the number of rows is $B-1$ and the number of parameters, δ_i's (i=1 0, ..., B–1 0, w_0), is B from each p_{iB} row of matrix **D**. Thus, it provides a ratio of δ_i to δ_j. The choice of i and j is arbitrary, so let

i=(B–1 0) where $B-1 \in S_{k^*}$ for $0<k<K$ and $j=w_0$. With that, we get the system of three (=2+1 or B_{k^*}+1) equations which are composed of three δ_i's ($i \in S_{k^*}$, w_k for $0<k<K$) from each p_{ij} row (i, $j \in S_{k^*}$ for $0<k<K$). As each partial differential quotient is different, $\delta_{i0}=0$ for all i ($i \in S_{k^*}$) and $\delta_{w_k}=0$ for all k ($0<k<K$). Finally, for all the parameters, it works out that

$$\delta_{10} = \cdots = \delta_{B-1,0} = \delta_{w_0} = \delta_{w_{K-1}} = 0, \tag{a3}$$

showing that linear independence of each vector is evident. In the same way, we can ascertain this for the case of $B_{k^*}>2$. However, if $B_{k^*}=1$, then a certain relation between δ_{i0} ($i \in S_{k^*}$) and δ_{w_k} with respect to a submarket S_k ($0<k<K$) exists and it turns out that each vector is linearly dependent. Consequently, MIGHT 3 is identifiable in the case where the number of *SkS* brands for the submarket S_K is one and that for a submarket S_k is not less than two. This condition can be included in the first condition because of the arbitrariness of the choice of a submarket S_k, or we can consider k as K arbitrarily. Thus, given that the number of parameters is not greater than [B×(B+1)/2]–1, *MIGHT 3 is always identifiable in the case where the number of SkS brands of a submarket is greater than or equal to two.*

Identifiability of MIGHT 2

The matrix **D** for MIGHT 2 is composed of 2B–1 columns. We denote the column vector of the first-order partial derivative as $\nabla^{ik}=\{\partial p_{ij} / \partial \pi_{ik}\}$ and $\nabla^{w_k}=\{\partial p_{ij} / \partial w_k\}$, each of which has [$B$×($B$+1)/2]–1 rows. The necessary and sufficient condition for the identifiability is that each vector is linearly independent or that there exists no non-trivial solution to the following equation:

$$\begin{aligned}\delta_{10}\nabla^{10} + \cdots + \delta_{B-1,0}\nabla^{B-1,0} + \delta_{1,1}\nabla^{1,1} + \cdots + \delta_{B_1-1,1}\nabla^{B_1-1,1} + \cdots \\ + \delta_{ik}\nabla^{ik} + \cdots + \delta_{B_{K-1},K}\nabla^{B_{K-1},K} + \delta_{w_0}\nabla^{w_0} + \cdots + \delta_{w_{K-1}}\nabla^{w_{K-1}} = 0\end{aligned} \tag{a4}$$

Regarding each p_{iB} row of matrix **D**, we can get an equation system where the number of rows is B–1 and B is the number of parameters, δ_i's (i=1 0, ..., B–1 0, w_0). Hence, a rational relation of δ_i and δ_j is procured. We can choose i and j arbitrarily, so let i=(1 0) and $j=w_0$. We consider the two cases, where $B>B_1+1$ and where $B=B_1+1$, separately.

In the case where $B>B_1+1$, taking notice of that relation, each p_{ij} (i=1, ..., B_1; $j=B_1+1$, ..., B–1) row gives the system of B_1×(B–B_1–1) equations of which the parameters are $B-B_1-1$ δ_{i0}'s ($i \in S_k$ for $k\geq 2$) and δ_{w_0}. It turns out that $\delta_{i0}=0$ for all $i \in S_k$ ($k\geq 2$) and $\delta_{w_0}=0$. Next, reducing the equation systems for each submarket S_k ($0<k<K$) individually, we acquire the system of B_k×(B_k+1)/2 equations of which the parameters are B_k–1 δ_{ik}'s ($i \in S_k$) and δ_{w_k} for

each submarket S_k $(0<k<K)$. Consequently $\delta_{ik}=\delta_{w_k}=0$ $(i\in S_k)$ is derived for each submarket S_k $(0<k<K)$ separately. Finally, from each p_{ij} $(i,\ j\in S_k)$ row, we procure the system of $B_K\times(B_K+1)/2$ equations of which the parameters are B_K-1 δ_{iK}'s $(i\in S_k)$. Accordingly, we get $\delta_{iK}=0$ $(i\in S_k)$. As a result, the above discussion brings

$$\delta_{10}=\cdots=\delta_{B-1,0}=\delta_{1,1}=\cdots=\delta_{B_1-1,1}=\cdots=\delta_{ik}=\cdots=\delta_{B_K-1,K}=\delta_{w_0}=\cdots=\delta_{w_{K-1}}=0 \tag{a5}$$

Next, we consider the case where $B=B_1+1$. This is exactly the case where $K=2$ and $B_2=1$. Each of the p_{1j} $(j=1, ..., B-1)$ rows gives the system of B–1 equations which includes $2B-2$ $(=B-1+B-2+1)$ parameters. Hence, we get the rational relations among $B-1$ parameters. Because their choice is arbitrary, we regard them as δ_{i0} $(i=1, ..., B-1)$. Likewise, we obtain the system of $B-2$ equations which includes $2B-3$ $(=B-2+B-2+1)$ parameters from each p_{2j} $(j=2, ..., B-1)$ row. Thus, the rational relation among $B-1$ parameters δ_{i0} $(i=1, ..., B-1)$ is secured. Similarly we can get the rational relations among $B-1$ parameters δ_{i0} $(i=1, ..., B-1)$ from each p_{ij} $(i=3, ..., B-1; j=i, ..., B-1)$. These operations produce the system of independent $B-1$ equations, where the parameters are each of δ_{i0} $(i=1, ..., B-1)$. Consequently it turns out that $\delta_{10}=\cdots=\delta_{B-1,0}=\delta_{w_0}=0$. Next, we can derive the system of $B-1$ equations of which the parameters are $B-2$ δ_{i1} $(i=1, ..., B-2)$ and δ_{w_1} from each $p_{i,B-1}$ $(i=1, ..., B-1)$. Accordingly, it gives $\delta_{11}=\cdots=\delta_{B-1,1}=\delta_{w_1}=0$. Eventually we get

$$\delta_{10}=\cdots=\delta_{B-1,0}=\delta_{1,1}=\cdots=\delta_{B_1-1,1}=\cdots=\delta_{ik}=\cdots=\delta_{B_K-1,K}=\delta_{w_0}=\cdots=\delta_{w_{K-1}}=0 \tag{a6}$$

The case $B=2$ is the only one which does not satisfy the necessary condition $2B-1\leq[B\times(B+1)/2]-1$. The assumption that the last brand B is the *SkS* brand of the submarket S_K is held for all cases in MIGHT 2. Thus, *MIGHT 2 is always identifiable if the number of brands is greater than two.*

Identifiability of MIGHT 1

The number of columns of matrix **D** for MIGHT 1 is B. We indicate the column vector of the first-order partial derivative as $\nabla^i=\{\partial p_{ij}\ /\ \partial\pi_{i0}\}$ and $\nabla^{w_0}=\{\partial p_{ij}\ /\ \partial w_0\}$ of which the dimension is $[B\times(B+1)/2]-1$. The necessary and sufficient condition for the identifiability is that each vector is linearly independent or that there exists no non-trivial solution to the following equation:

$$\delta_1 \nabla^1 + \cdots + \delta_{B-1,0} \nabla^{B-1,0} + \delta_{w_0} \nabla^{w0} = 0 \qquad \text{(a7)}$$

With respect to each p_{iB} row of matrix **D**, we can get the equation system in which the number of rows is B–1 and the number of parameters, δ_i's (i=1, ..., B–1, w_0), is B. Thus, it gives a ratio of δ_i to δ_j. Since their choice is arbitrary, let i=1 and $j=w_0$. On the ground that brand 1 always belongs only to the submarket S_1 and that the differential quotients are different from each other, we gain $\delta_1=\delta_{w_0}=0$ from the p_{11} row. In the same way, from each p_{ii} row we can derive $\delta_i = 0$. Finally, it turns out

$$\delta_1 = \cdots = \delta_{B-1,0} = \delta_{w_0} = 0 \qquad \text{(a8)}$$

The presumption that the last brand B is the *SkS* brand of the submarket S_K is always assured in MIGHT 1. In addition, all the cases in MIGHT 1 promise the necessary condition, $B \le [B \times (B+1)/2] - 1$. Consequently *MIGHT 1 is identifiable for any case.*

References

Bass, Frank M., Moshe M. Givon, Manohar U. Kalwani, David Reibstein, and Gordon P. Wright (1984), "An Investigation into the Order of the Brand Choice Process," *Marketing Science*, **3, 4** (Fall), 267-87.

Bentler, Peter M. (1980), "Multivariate Analysis with Latent Variables: Causal Modeling," *Annual Review of Psychology*, **31**, 419-56.

Bozdogan, Hamparsum (1987), "Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions," *Psychometrika*, **52, 3** (September), 345-70.

Bourgeois, Jacques C., George H. Haines, Jr., and Montrose S. Sommers (1987), "Product Market Structure: Problems, Definitions, and Issues," In Michael J. Houston (eds.), *Review of Marketing*. AMA, 327-84.

Clogg, C. D. (1977) , "Unrestricted and Restricted Maximum Likelihood Latent Structure Analysis: A Manual for Users," working paper #9, Population Issues Research Office, The Pennsylvania University.

Colombo, Richard A. and Donald G. Morrison (1989), "A Brand Switching Model with Implications for Marketing Strategies," *Marketing Science*, **8, 1** (Winter), 89-99.

Cooper, Lee G. and Masao Nakanishi (1988), *Market Share Analysis: Evaluating Competitive Marketing Effectiveness*. Boston: Kluwer Academic Publishers.

Cooper, Lee G. and Akihiro Inoue (1996), "Building Market Structures from Consumer Preferences," *Journal of Marketing Research*, **33** (August), 293-306.

Day, George, Allan D. Shocker, and Rajendra K. Srivastava (1979), "Customer-Oriented Approaches to Identifying Product Markets," *Journal of Marketing*, **43** (Fall), 8-19.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), "Maximum Likelihood from Incomplete Data Via the EM Algorithm," Journal of the Royal Statistical Society, B 39, 1-38.

Dillon, William R., Ajith Kumar, and Melinda Smith de Borrero (1993), "Capturing Individual Differences in Paired Comparisons: An Extended BTL Model Incorporating Descriptor Variables," *Journal of Marketing Research*, **30** (February), 42-51.

Efron, Bradley (1979), "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, **7, 1,** 1-26.

Efron, Bradley and Robert J. Tibshirani (1993), *An Introduction to Bootstrap*. New York, Chapman & Hall, Inc.

Goodman, Leo A. (1974), "Exploratory Latent Structure Analysis Using both Identifiable and Unidentifiable Models," *Biometrika*, **61**, 215-31.

Grover, Rajiv and William R. Dillon (1985), "A Probabilistic Model for Testing Hypothesized Hierarchical Market Structures," *Marketing Science*, **4, 4** (Fall), 21-40.

Grover, Rajiv and V. Srinivasan (1987), "A Simultaneous Approach to Market Segmentation and Market Structuring," *Journal of Marketing Research*, **24** (May), 139-53.

Harshman, Richard A., Paul E. Green, Yoram Wind, and Margaret E. Lundy (1982), "A Model for the Analysis of Asymmetric Data in Marketing Research," *Marketing Science*, **1** (Spring), 205-42.

Hauser, John R. and Birger Wernerfelt (1990), "An Evaluation Cost model of Consideration Sets," *Journal of Consumer Research*, **16** (March), 393-408.

Inoue, Akihiro and Masao Nakanishi (1991), "An Extension of the PRODEGY model - An Application of Latent-Class Analysis to Forced Switching Data," working paper, #003, School of Business Administration, Kwansei Gakuin University.

Inoue, Akihiro (1993), "A System of Models to Identify Competitive Groups with Population Heterogeneity: MIGHT System," Research paper, John E. Anderson Graduate School of Management, University of California, Los Angeles.

Jain, Dipak, Frank M. Bass, and Yu-Min Chen (1990), "Estimation of Latent Class Models with Heterogeneous Choice Probabilities: An Application to Market Structuring," *Journal of Marketing Research*, **27** (February), 94-101.

Jedidi, Kamel, Venkatram Ramaswamy, and Wayne S. DeSarbo (1993), "A Maximum Likelihood Method for Latent Class Regression Involving A Censored Dependent Variable," *Psychometrika*, **58, 3**, 375-94.

Kalwani, Manohar U. and Donald G. Morrison (1977), "A Parsimonious Description of the Hendry System," *Management Science*, **23** (January), 467-77.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, **26** (November), 379-90.

Kumar, Ajith and C. M. Sashi (1989), "Confirmatory Analysis of Aggregate Hierarchical Market Structures: Inferences from Brand-Switching Behavior," *Journal of Marketing Research*, **26** (November), 444-53.

McHugh, Richard B. (1956), "Efficient Estimation and Local Identification in Latent Class Analysis," *Psychometrika*, **21, 4** (December), 331-47.

McLachlan, Geoffrey and Kaye E. Basford (1988), *Mixture Models: Inference and Applications to Clustering*. N. Y.: Marcel Dekker Inc.

Mooijaart, Ab and Peter G. M. van der Heijden (1992), "The EM Algorithm for Latent Class Analysis with Equality Constraints," *Psychometrika*, **57, 2** (June), 261-69.

Novak, Thomas P. and Charles Stangor (1987), "Testing Competitive Market Structures: An Application of Weighted Least Squares Methodology to Brand Switching Data," *Marketing Science*, **6, 1** (Winter), 82-97.

Novak, Thomas P. (1993), "Log-Linear Trees: Models of Market Structure in Brand Switching Data," *Journal of Marketing Research*, **30** (August), 267-87.

Rao, Vithala R. and Darius J. Sabavala (1981), "Inference of Hierarchical Choice Processes from Panel Data," *Journal of Consumer Research*, **8** (June), 85-96.

Sabavala, Darius J. and Donald G. Morrison (1977), "A Model of TV Show Loyalty," *Journal of Advertising Research*, **17** (December), 35-43.

Siddarth, S., Randolph E., Bucklin, and Donald G. Morrison (1993), "Making the Cut: Modeling and Analyzing Choice Set Restriction in Scanner Panel Data," working paper, John E. Anderson Graduate School of Management, UCLA.

Shocker, Allan D., David W. Stewart, and Anthony J. Zahorik (1990), "Market Structure Analysis: Practice, Problems, and Promise," In George Day, Barton Weitz, and Robin Wensley (eds.), *The Interface of Marketing and Strategy*, JAI Press Inc., 9-56.

Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara, and Prakash Nedungadi (1991), "Consideration Set Influences on Consumer Decision Making and Choice; Issues Models and Suggestions," *Marketing Letters*, **2**, 181-98.

Titterington, D. M., A. F. M. Smith, and U. E. Makov (1985), *Statistical Analysis of Finite Mixture Distributions*. N. Y.: John Wiley and Sons.

Urban, Glen L. , Philip L. Johnson, and John R. Hauser (1984), "Testing Competitive Market Structures," *Marketing Science*, **3, 2** (Spring), 83-112.

Winsberg, Suzanne and Geert de Soete (1993), "A Latent Class Approach to Fitting the Weighted Euclidean Model, CLASCAL," *Psychometrika*, **58, 2** (June), 315-30.

Zenor, Michael J. and Rajendra K. Srivastava (1993), "Inferring Market Structure With Aggregate Data: A Latent Segment Logit Approach," *Journal of Marketing Research*, **30** (August), 369-79.