한국경영과학회/대한산업공학회 '97 춘계공동학술대회. 1997년 4월 25일-26일. 포항공과대학교

Efficiency of Direct Estimation Method in Batsell and Polking's Model

Sang-June Park*, Minhi Hahn**, Changhoon Shin***

* Doctoral Candidate, Graduate School of Management, KAIST
** Professor, Graduate School of Management, KAIST
*** Assistant Professor, Dept. of Logistics Engineering, Korea Maritime University

Abstract

The mathematical models such as Tversky's Elimination-by-Aspects model (EBA), Tree models, and PROBIT model have been developed to get over the IIA (Independence from Irrelevant Alternatives) problem found in the Luce-type choice models. However, these models have some limitations due to the assumption of tree structure and estimation method. Batsell and Polking (Marketing Science 1985; hereafter BP) proposed a new model that overcomes such limitations of the models. The BP model gives us managerial insight into the sources of competition and helps us diagnose the extent to which the observed market shares deviate from those estimated by the Luce-type choice models.

One of the notable characteristics of BP model is that parameters can be estimated indirectly by the ordinary least squares (OLS). Although BP claim that it is convenient to indirectly estimate parameters through intermediate parameters, we find that it is more efficient and informative to directly estimate the parameters in several ways. In this article, we present a direct method in detail and show that the direct method provides more benefits than the indirect method. Three of these benefits are discussed as follows:

First, we can divide the original estimation problem into several sub-problems to estimate parameters. Because of the divisibility, the direct method is easier to estimate parameters than the indirect method.

Second, when we want to focus on the competition between our product and a particular competitor (or competitors), direct method needs the market-share data of the choice sets which only contain both our product and the competitor (or competitors) of interest. In contrast, the indirect method needs the market-share data for all the possible choice sets. One of limitations of BP model is that it needs too much efforts to gather the input data. The divisibility property of the direct method can decreases such efforts.

Third, the degree of deviation from the IIA property can be tested for any pair of products in the choice sets when the direct method is used. But the same does not apply to the indirect method.

1. Introduction

After Debreu (1960) first noted the IIA problem of the Luce-type choice models, EBA model (Tversky, 1972), Tree models (Tversky and Sattath,1979; McFadden, 1981), and PROBIT model (e.g., Currim, 1982) have been developed to overcome the problem. However, the Tree models require a priori specification of the aspects representing choice alternatives and tree structure of these aspects. On the other hand, EBA and PROBIT have difficulty in estimating parameters. Batsell and Polking (1985) offered a theorem and proof which guaranteed the existence of a new model that overcomes the problems of the previous models. They phrased that information of market-shares can be transformed into the competitive effect which competing products have on each other's market-share. The mathematical model proposed by BP is widely accepted as one of important choice models that can overcome the IIA problem.

BP model can yield insight into the underlying competitive structure that suggest relevant attributes and potential tree structures (Batsell and Polking, 1985). BP model also has potential to test the competitive structure when the product attributes governing choice in the market are not well known, or the hierarchy of the tree is not clear a priori. Well-known testing models of competitive market structures (Kannan and Wright, 1991; Novak and Stangor, 1987; Grover and Dillon, 1985; Urban, Johnson, and Hauser, 1984) typically require the assumption of tree structure a priori.

Gensch and Ghose (1992) combined a tree model and the BP model into EBD model (Elimination by Dimensions model). The EBD showed much better fits than a tree model or a logit model.

However, BP model has not enjoyed widespread application in the field of marketing because it takes much efforts to collect data (Currim, Meyer, and Le, 1988). We find that the limitation of their model can be overcome to some extent by using the direct method which we discuss in this paper. Additionally, although BP claim that it is convenient to estimate parameters indirectly through intermediate parameters, we find that the direct method is more efficient and informative than the indirect method in several ways.

2. Direct Method vs. Indirect Method

In this article, we investigate the indirect as well as the direct method. Especially, we discuss the direct method in detail because BP hardly explained this method.

2.1. Direct Method

Basically, we use the same notations as used by BP. Let T={1, 2, ... , N} be a set of alternative products with which we want to analyze the competitive effects. For each subset $A \subset T$, let $P_A(i)$ denotes the probability that an alternative i will be chosen when the set A is a set of available alternatives. BP assumed no a priori relationship between the probabilities P_A's. They only assumed that $P_A(i) \neq 0$ for all $i \in A$ and for all $A \subset T$.

For each non-empty subset $A \subset T$ and for i, $j \in A$, they define

$$\beta_{ij}^{A} = \ln\left(\frac{P_A(i)}{P_A(j)}\right). \qquad (1)$$

When the ratio of i's share to j's share is defined as equation (1), they proved that there exist unique numbers α_{ij}^{I} defined for i, $j \in T$, $I \subset T$ with $I \cap \{i, j\} = \varnothing$ such that for every subset $A \subset T$,

$$\beta_{ij}^{A} = \sum_{I \subset [A-\{i,j\}]} \alpha_{ij}^{I}. \qquad (2)$$

They interpreted that α_{ij}^{I} measures the effect of presence of the products in the set I on the competition between i and j. They called α_{ij}^{I} an nth order effect when the number of products is n-2 in the set I (i.e., n = #($I \cup \{i, j\}$), and defined equation (2) as the nth order model when the effects of the variables with the order higher than n do not exist ($\alpha_{ij}^{I} = 0$ if # (I) > (n-2)). With this definition, they interpreted the Luce model as 2[nd] order model and they insisted that the model is new when the order is greater than 3. They called their model 'the saturated model' when the order equals the number of alternatives.

Saturated Model: BP showed that α_{ij}^{I} could be expressed as the equation (3).

$$\alpha_{ij}^{I} = (-1)^{\#(I)} \sum_{J \subset I} (-1)^{\#(J)} \beta_{ij}^{ijJ} \qquad (3)$$

For the sake of convenience, we rewrite equation (1) as equation (4).

$$\beta_{ij}^{J} = \ln\left(\frac{P_A(i)}{P_A(j)}\right) \quad \text{for } J=A-\{i,j\} \qquad (4)$$

And equation (2) is rewritten as equation (5).

$$\beta_{ij}^{J_m} = \sum_{I \subset J_m} \alpha_{ij}^{I}, \ m=1, 2, ..., M \qquad (5)$$

$$\text{where, } M = \sum_{r=1}^{n-2} {}_{n-2}C_r = 2^{n-2}$$

For the alternatives (i, j), when we observe all possible sets $\{J_m, i, j\}$, we obtain as many as M simultaneous equations (M: the number of subsets of J). We can obtain the equation (6) or (7) if $\beta_{ij}^{J_m}$'s and $\alpha_{ij}^{J_m}$'s are arranged in the order of # (J_m).

$$\begin{bmatrix} \beta_{ij}^{J_1} \\ \beta_{ij}^{J_2} \\ \cdot \\ \beta_{ij}^{J_m} \end{bmatrix} = \begin{bmatrix} 1 & 0 & . & 0 \\ 1 & 1 & . & 0 \\ . & . & . & . \\ 1 & 1 & . & 1 \end{bmatrix} \begin{bmatrix} \alpha_{ij}^{J_1} \\ \alpha_{ij}^{J_2} \\ \cdot \\ \alpha_{ij}^{J_m} \end{bmatrix} \qquad (6)$$

$$B^{ij} = C^{ij} A^{ij} \qquad (7)$$

$$\text{where, } B^{ij} = \left[\beta_{ij}^{J_1} \ \ \beta_{ij}^{J_2} \ \ . \ \ \beta_{ij}^{J_m}\right]',$$

$$A^{ij} = \left[\alpha_{ij}^{J_1} \ \ \alpha_{ij}^{J_2} \ \ . \ \ \alpha_{ij}^{J_m}\right]'$$

In equation (7), we find the matrix C^{ij} has some particular properties which make it possible to solve the equation efficiently.

Properties of $C^{ij} = \left[c_{m_1 m_2}\right]$:

(1) $c_{m_1} = 1$, for $m_1 = 1, ..., M$

(2) $c_{M m_2} = 1$, for $m_2 = 1, ..., M$

(3) $c_{m_1 m_2} = 1$, for $m_1 = m_2$

(4) $C^{ij} = C$ (all i,j; i < j)

Property (4) means that matrix C^{ij} remains unchanged even though B^{ij} and A^{ij} change according to product i and j. When we wish to obtain all possible A^{ij} (all i,j; i < j), we can aggregate the equation (7), so that we make an aggregated matrix where the sub-vectors B^{ij} and A^{ij} are arranged first in the order of i and then in the order of j again as equation (8).

$$\begin{bmatrix} B^{12} \\ B^{13} \\ \cdot \\ B^{n-1n} \end{bmatrix} = \begin{bmatrix} C & 0 & \cdot & 0 \\ 0 & C & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & C \end{bmatrix} \begin{bmatrix} A^{12} \\ A^{13} \\ \cdot \\ A^{n-1n} \end{bmatrix} \quad (8)$$

or as

$$B = EA \quad (9)$$

We can obtain the following unique solution for equation (9).

$$A = \begin{bmatrix} A^{12} & A^{13} & \cdot & A^{n-1n} \end{bmatrix}' = E^{-1}B$$

The solutions for equation (9) can be obtained from equation (7) because off-diagonals of E matrix are zero matrices. Furthermore, in each sub-problem (equation (7)), the matrix C represents the lower-triangular matrix and all diagonal elements in C are 1. Thus, C is a nonsingular matrix (See the properties, (1), (2), and (3)). So we can always obtain the following unique solution.

$$A^{ij} = \begin{bmatrix} a_{ij}^{J_1} & a_{ij}^{J_2} & \cdot & a_{ij}^{J_m} \end{bmatrix}' \text{ as } A^{ij} = C^{-1}B^{ij}$$

Estimation model: The lower order versions of equation (9) can be represented by equation (10) or (11). In these equations, vectors A_R and A_R^{ij} can be obtained after eliminating higher order solutions from the vector A and A^{ij}, respectively. Matrix E_R and C_R represent the relationship between the solution vectors and log-transformed ratios of market-shares in the original problem and sub-problem, respectively. The following equation (10) gives us the OLS estimators corresponding to the original problem, whereas equation (11) gives us the OLS estimators corresponding to the sub-problems.

(Original problem) $B = E_R A_R + e$ (10)

(Sub-problem) $B^{ij} = C_R A_R^{ij} + e^{ij}$ (all i,j; i < j) (11)

Estimators of the original problem can be obtained by estimating parameters of the sub-problems, because the diagonals of matrix E_R are C_R matrices and the off-diagonals of E_R are zero matrices.

2.2 *Indirect method*

Now, we will briefly review the indirect method described by BP. They found that α_{ij}^{I}'s can be expressed as the linear combination of s_i^I's as

$$\alpha_{ij}^{I} = s_i^{iI} + s_i^{ijI} - (s_j^{jI} + s_j^{ijI}) \quad (12)$$
$$\text{if } I \cap \{i,j\} = \varnothing .$$

They proposed a method such that the α_{ij}^{I}'s are indirectly calculated through an equivalent set of scales s_i^I's with equation (12). To compute the s_i^I's, they derived the equation (13) but the scales are not linearly independent, so that we can not obtain the scales without independent condition of the scales. They derived the equation (14) satisfying the linearly independent condition (BP termed it as normalizing equations) to guarantee uniqueness of the scales.

$$\ln\left(\frac{P_A(i)}{\bar{P}_A}\right) = \left(\frac{N_A - 1}{N_A}\right)\left(\sum_{I \subset A} s_i^I\right) + \left(-\frac{1}{N_A}\right)\left(\sum_{j \in A, j \neq i} \sum_{I \subset A} s_j^I\right) \quad (13)$$

$$U_I = \sum_{i \notin I} s_i^{iI} \quad (14)$$

where, $\bar{P}_A$ is geometric mean of $P_A(i)$'s.
N_A is the number of products in choice set A.

There are two ways of obtaining the scales s_i^I's which satisfy equation (13) and (14). One way is to combine equation (13) and (14). The other is to substitute the linearly dependent scales with linearly independent scales. BP measured scales (s_i^I's) with the former way.

2.3 *Tests of IIA in Both Methods*

With the indirect method, BP conducted the statistical test (F-test) of relative significance of different order models. With the test, BP diagnosed whether the higher order model is different from the 2nd order model (equivalent to the Luce model). However, the order of the indirect method does not correspond to the order of the direct method (See equation (12)). For example, the particular intermediate variables (i.e., s_i^I satisfying #(I) = r) are related to the α_{ij}^{I} with r and r+1 order effects simultaneously. Thus, we can not interpret the test of indirect method as that of the relative significance of different order models for α_{ij}^{l}, even though the tests of indirect method can diagnose the IIA problem.

With the direct method, we can show that the test can be done with the sub-problems as well as with the original problem. In equation (10), the error sum-of squares (ESS) can be obtained from equation (15).

$$e'e = \sum_{i<j} e^{ij'} e^{ij} \tag{15}$$

where, $e = B - E_R A_R$
$e^{ij} = B^{ij} - C_R A_R^{ij}$

The right-hand side of equation (15) is the ESS of the original problem, whereas the left-hand side of the equation is the sum of ESS of each sub-problem (See equation (11)). After calculating the ESS of each sub-problem without ESS of the original problem, we can obtain R^2 and conduct F-tests on the each sub-problem as well as on the original problem.

3. Conclusion

Although BP claim that it is convenient to estimate parameters indirectly through intermediate parameters, we find that the direct method is more efficient and informative than the indirect method in the following aspects.

Efficiency in parameter estimation: There are fewer parameters to estimate in the indirect method. However, the number of parameters to estimate increases as the number of the alternatives increases. In this case, the direct method can be more efficient because we can divide the original problem into several sub-problems. Then we can estimate parameters of each sub-problem. The efficiency in estimation can be obtained in the calculation of the inverse matrix.

Efficiency in data collection: One limitation of BP model is that it needs much efforts to gather the input data. The divisibility property of the direct method decreases the efforts in the following case. Assume that marketing practitioners want to focus on the direct competition between their product and a particular competitor. For example, when they only focus on the competition between their product i and a competitor's product j. In the direct method, they do not have to observe market-shares in all the possible choice sets. They need to observe market-shares in each subset which only contains the product i and j. In contrast, the indirect method needs the market-share data in all the possible choice sets.

More information on competition: The direct method gives us information (R^2 and ESS) on each sub-problem which the indirect method does not give. Thus, in the direct method, test of IIA property can be applied to any pair of products in the choice sets. The same is not true of indirect method.

Accurate testing of IIA property: The test of the indirect method does not give us the relative significance of different order models directly related to α_{ij}^{l}. Thus we need the test of the direct method to identify the relative significance of different order models for α_{ij}^{l}.

References

[1] Batsell, Richard R. and Polking (1985), "A New Class of Market Share Models," *Marketing Science*, 4 (Summer), 177-98.

[2] Currim, I. S. (1982), "Predictive Testing of Consumer Choice Models Not Subject to Independence of Irrelevant Alternatives," *Journal of Marketing Research*, 19 (May), 208-222.

[3] _______ , Robert J. Meyer, and Nhan T. Le (1988), "Disaggregate Tree Structured Modeling of Consumer Choice Data," *Journal of Marketing Research*, Vol. 25 (August), 301-309.

[4] Debreu, G. (1960), "A Review of Individual Choice Behavior: A Theoretical Analysis," *American Economic Review*, 50 (December), 186-88.

[5] Gensch, P. E. and Sanjoy Ghose (1992), "Elimination by Dimensions," *Journal of Marketing Research*, Vol. 26 (November), 417-29

[6] Grover, Rajiv and William R. Dillon (1985), "A Probabilistic Model for Testing Hypothesized Hierarchical Market Structures," *Marketing Science*, Vol. 4, no. 4 (Fall), 312-335.

[7] Kannan, P. K. and G. P. Wright (1991), "On Testing Competitive Market Structures," *Marketing Science*, Vol. 10, no. 4 (Fall).

[8] McFadden, Daniel (1981), "Econometric Models of Probabilistic Choice," in Structural Analysis of Discrete Data with Economic Applications, C.F. Manski and D. McFadden (eds.), Cambridge: MIT Press.

[9] Novak, Thomas P. and Charles Stangor (1987), "Testing Competitive Market Structures: An Application of Weighted Least Squares Methodology to Brand Switching Data," *Marketing Science*, Vol. 6, no. 1 (Winter), 82-97.

[10] Tversky, Amos (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79, 281-99.

[11] _______ and Shmuel Sattath (1979), "Preference Trees," *Psychological Review*, 86 (November), 542-93.

[12] Urban, G., P. L. Johnson, and J. R. Hauser (1984), "Testing Competitive Market Structures," *Marketing Science*, Vol. 3, no. 2 (Spring), 83-112.