A Probabilistic Choice Model of Multiple Items Selection

著者	Ogawa Kohsuke, Kawabata Hiroyuki, Moriguchi Takeshi, Uchida Manabu
出版者	法政大学産業情報センター
journal or publication title	グノーシス : 法政大学産業情報センター紀要
volume	4
page range	59-79
year	1995-03-31
URL	http://hdl.handle.net/10114/00020879

A Probabilistic Choice Model of Multiple Items Selection +

Kohsuke Ogawa
Hiroyuki Kawabata
Takeshi Moriguchi
Manabu Uchida

ABSTRACT

Conventional choice model of consumer behavior posits that buyers choose a single item or one brand at a time. However, there are many occasions that consumers pick up several items or a couple of brands at the same time to satisfy their own needs and/or to maximize their family's utility. We can observe this kind of phenomena in the market of packaged consumer products such as tobacco, candy bar, beer, and soft drink.

We construct a new probabilistic choice model of consumer behavior. In order to deal with such a situation that a consumer makes a simultaneous selection of multiple items from his consideration set of brands, we assume a two stage model of consumer choice behavior. At first stage, we suppose, a consumer makes a decision whether he buys a single brand or a mixed bundle of brands. Then he makes a decision of which brand(s) to be picked up and assigns the allocation number of multiple brands.

We use a set of point-of-sale beer data scanned at a convenience store to estimate the model parameters. Our model is well fitted into the empirical data. Marketing implications and the possibilities of further extensions with our basic idea are provided.

1. Introduction

1-1 Brand Choice Models

It has already passed about twenty five years after the first generation of marketing scientists read through a milestone seminar book "Stochastic Models of Buying Behavior" by 3M (Massy, Montgomery, Morrison) in 1970. Since then, a lot of works have been done to extend their original idea to three directions.

+ This paper was originally presented at the annual meeting of the Asian Pacific Operations Research Society of 1994 which was held at Fukuoka, Japan, on July 27, 1994

First one is to make the choice model as simple as possible. Based upon a few basic distributional assumptions of consumer choice probabilities, British scholars, e.g., Ehrenberg, Goodhardt, Uncle, have tried to explain the market phenomena and/or market trend such as development of penetration rate, brand switching, and brand loyalty, more clearly (Ehrenberg 1988). (*1) They tend to like an elegant theory and a rather parsimonious way in modeling efforts.

Second one is the approach pursued by American professors. Most of them are scholars in the second generation who have been influenced by one of 3M, i.e. Professor Donald Morrison. Among them, for example, are Schmittlein, Colombo, Fader, and so forth. (*2) Of interest is that second generation of brand choice marketing scientist in the United States have been consistently focusing upon the empirical findings. Perhaps it is partly because of American tradition of pragmatism, and partly because scanner data (POS data) has been made readily available for marketing applications (Guadagni and Little 1982). (*3)

Third one is the group of "variety-seeking" modelers who are influenced by consumer behaviorists who come from the academic department of psychology or sociology, although 3M has a definite effect on this cohort.

Kahn, McAlister, Givon are among others who are interested in the occasions that people tend to switch from one to another brand in order to avoid boring, or in some cases to seek for a stimulation. They could account for a broad variety of switching behaviors, empirically (McAlister 1979) and theoretically (Kahn, Kalwani, Morrison 1986). (*4)

1-2 Multiple Items Selection

All probabilistic brand choice models above mentioned have a common modeling assumption. That is, a consumer buys one brand at a time. Very often the probability with which he chooses a brand is supposed to be independent from time to time (zero-order assumption). Or it might depend on the last purchase (first-order Markov assumption). Or model builders suppose that the odds should be determined by the entire purchase history of consumer's brand choice (learning assumption). (*5)

Of course, brand choice theory could account for a choice situation when a consumer chooses multiple packs of diapers of the same brand (Nakajima 1993). (*6) However, any conventional models cannot deal with such an occasion that she picks up a package of P&G's New Pampers for night use, and a pack of Kao's New Merries for regular use at the same time.

Modeling convention is that, they suppose, she might locate a New Pampers at first and then switch to a New Merries, and vice versa. Alternatively, since they can not identify either one of possibilities, i.e., a switch from Pampers to Merries, or from Merries to Pampers, they could "blend" two possibilities with a chance of 50-50 percent.

In any sense, the above treatment is not a theoretically-sound procedure, but just a conventional wisdom to get around some (not many) irregularity of purchase events. We would say that a much more rigorous modeling effort should be made to cope with such a convention.

Thus in this article, we will attempt to dealing with a choice situation in that a consumer picks up more than two items of different brands at one time. A formal choice model of multiple items will be constructed by assuming a two-stage process model.

1-3 Outline of this Article

In the next section, we begin with a formal definition of our new probabilistic choice model in that a consumer makes a simultaneous selection of multiple items. A two-stage model will be presented.

In section 3, we describe a special feature of beer purchase records. We make use of a set of point-of-sale beer data scanned at a convenience store, then estimate the model parameters. Individual purchase records are aggregated into the choice shares by the number of items picked up by a customer at a time. We show that our model is well fitted into the empirical data.

Finally in section 4, marketing implications of our multiple items selection models are provided. And the possibilities of further extensions are discussed.

2. Model Descriptions

2-1 An Illustrative Example: A Beer Purchase Record

A short list of beer purchase records by a consumer is shown in Table 1. We can observe this kind of purchase behavior quite often when we keep track of beer purchase records at a liquor shop or a convenience store. As well, a similar pattern of consumer purchase records can be seen in such a product category as instant cup noodle, candy bar, or tobacco product.

The customer (#0841) bought eight cans of beer on April 7, 1991. Not surprisingly, he chose three different brands in two kinds of size, 350ml and 500ml. Note that there are three brands, "Ichiban-Shibori", "Beer-Ginjo", "Gin-Jikomi". All of three brands belong to the same sub-category, "thick in taste".

Although he purchased a single can of "Ichiban-Shibori" of 500ml size on May 12, he bought again eight cans of beer two weeks later, a single brand choice. They were chosen from three different brands, "Z", "Ichiban-Shibori", and "Gin-Jikomi". On the next occasion, he picked up twenty cans of "Ichiban-Shibori".

Table 1: Beer Purcahse History (Paneler = #0841)

Date	Brand	Maker	Size	Units
91/04/07	Ichiban-Shibori	Kirin	500 ml	4
	Beer-Ginjo	Suntory	500 mi	1
	Gin-Jikomi	Sapporo	350 ml	2
	Gin-Jikomi	Sapporo	500 ml	1
91/04/12	Ichiban-Shibori	Kirin	500 ml	1
91/05/26	Z	Asahi	500 ml	1
	Ichiban-Shibori	Kirin	500 ml	5
	Gin-Jikomi	Sapporo	500 ml	2
91/06/06	Ichiban-Shibori	Kirin	350 ml	12
	Ichiban-Shibori	Kirin	500 ml	8
91/06/16	Ichiban-Shibori	Kirin	500 ml	4
	Gin-Jikomi	Sapporo	350 ml	4
	Gin-Jikomi	Sapporo	500 ml	4
91/06/24	Ichiban-Shibori	Kirin	350 ml	4
	Ichiban-Shibori	Kirin	500 ml	4

2–2 Existence of Core Brand

It is very difficult for us to predict accurately what will happen to a consumer on the next purchase occasion at individual level. Which brand(s) will be picked up? How many units (cans) will be allocated to each brand?

However, we might identify a general rule of brand choice in case of multiple items selection for aggregation. Judging from the beer purchase records mentioned above, we can set up two basic assumptions regarding brand choice behavior when a consumer chooses more than two brands at a time. The first one will be discussed in this sub–section, and the second one will be dealt with in the following sub–section.

A consumer seems definitely to have his most favorite brand. For example, the customer #0841 in Table 1 chose "Ichiban–Shibori" on every occasion over his panel record. His choice probability of "Ichiban–Shibori" is quite high (more than 50%) and might be stable over time.

However, in most cases, he put some other brands, for instance, "Gin–Jikomi", "Z" in his shopping basket. Perhaps it is because he also wants to try other varieties of beer in taste. His choice probability for the most favorite brand reflects his intensity of brand loyalty.

In what follows, we call the most preferred brand "core brand". Accordingly, its zero–order choice probability is denoted by π_A. When we can estimate the parameter values, a comparison of those parameters can be made across brands as well as among customer groups.

2–3 Three Step Decision Process

Next we assume that a consumer goes through three steps when he makes a decision of multiple items selection. A flow chart of multiple brand choice model is depicted in Figure 1. (*7)

At first stage, he decides how many units (cans for beer) he would buy at a time. We might construct a kind of purchase incidence model used by Gupta (1991), or Schmittlein and Peterson (1994). (*8) However, our interest does not lie in a Poisson-type incidence model. Therefore, we directly begin with a model description from the next stage 2.

Given the number of units to buy, a consumer has to determine whether he purchases a single brand (Case I), or a mixed bundle of different labels (Case II). We define α to be a probability that he buys more than a couple of units of a single brand in a lump-sum. In practice, some consumers tend to buy a single sort of beer in a bundle, often in a six pack. Others tend to seek for a variety in taste, so with a mixed match of several brands.

The third step, if he proceeds to Case I, is to choose a brand in a bundle with a probability π_A. Otherwise, we suppose that he buys a bundle of another single brand. When he advances to Case II in the second stage, he must decide how many of the total units he assigns to his core brand.

Here we posit that the number of core brands chosen is distributed according to a binomial distribution with the zero-order choice probability π_A, which was defined in the sub-section 2-2. In other words, in Case II, a consumer tends to make an independent decision of core brand choice one by one with the same probability π_A in Case I. This assumption will be relaxed in the Adjusted Model later in sub-section 3-3, in order to adjust the probability premium of core brand in case of bundle purchase.

Figure 1: A Flow Chart of Multiple Items Choice Process

Step 1: To decide the number of units to buy

↓

Step 2: To decide which way to take as a next setp, to purchase some of a single brand (Case I), or to pick up a mixed bundle of labels (Case II) (with prob = α, bundle purchase parameter)

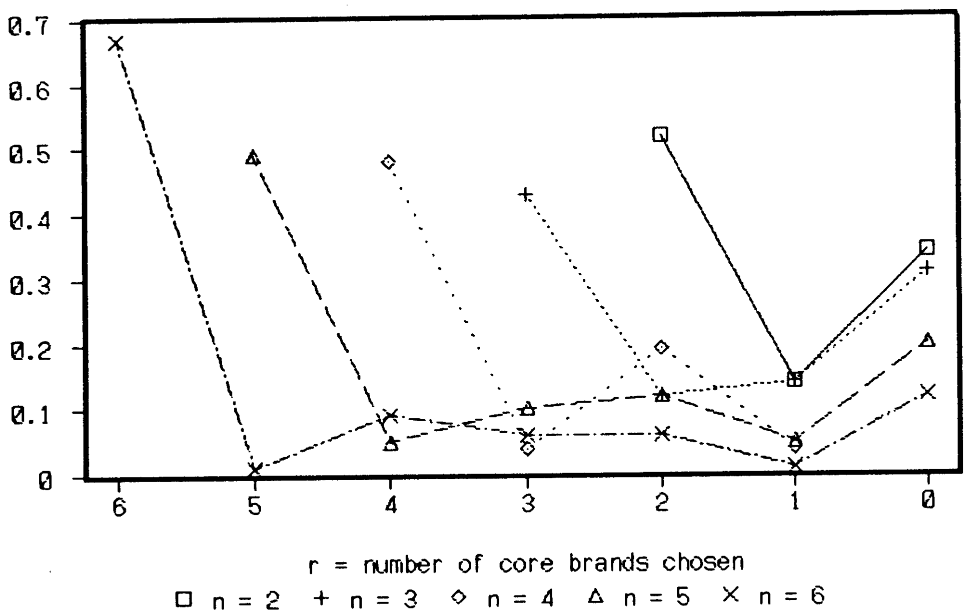
Prob. = α	**Prob. = $(1 - \alpha)$**
Step 3: Case I : Bundle Purchase To buy hiscore brand(with $\pi_A = P_0$), or the non-core brand(s) (with $\pi_A = 1 - P_0$)	Case II : Mixed Purchase to assign some of total units to his core brand, according to a binomial distribution (with $\pi_A = P_1$)

(*) In case of Basic Model or under U-shape Hypothesis, $\pi_A = P_0 = P_1$
in case of the Adjusted Model, π_A has to be adjusted to become $P_0 > P_1$.

Note that the last assumption (binomial distribution of allocated number of core brand) might be very plausible. It is supported by an empirical evidence below in the section 3.

One reason why we assume a binomial choice for core brand is that aggregate choice probabilities of core brand seem to be distributed according to a W-shaped pattern (Figure 2). This pattern can be explained by assuming a combination of two distributions: one with spikes in both ends (buying a single brand in a bundle) and the other with a binomial choice (buying a variety pack of several brands).

Figure 2: Probability Distributions: W-Shape of PA



2-4 Mathematical Formula for Brand Choice

Here we present a formal description of our brand choice model. We assume that a consumer faces a binary choice situation. That is, he has to choose between his core band and the "non-core" brand (a composite band of all other brands). For simplification, for the time being, we may omit a subscript (k) which stands for core brand.

Let the choice probability of core brand be π_A. Let n be the total number of units a consumer buys on each buying occasion, and r be the number of core brands which he picks up. Then the choice probability with which a consumer purchases r units of core brand becomes:

$$P(r) = \text{Prob}(\, r \mid n \,) \qquad \text{for } r = 0, 1, 2, \ldots, n \qquad (1)$$

Specifically,

$$P(n) = \alpha \, \pi_A + (1 - \alpha) \, \pi_A^{\,n} \qquad (1\text{-}n)$$

$$P(r) \doteq (1 - \alpha) \, {}_nC_r \, \pi_A^{\,r} \, (1 - \pi_A)^{(n-r)} \qquad (1\text{-}r)$$

$$P(0) \doteq \alpha \, (1 - \pi_A) + (1 - \alpha)(1 - \pi_A)^{n} \qquad (1\text{-}0)$$

Summing up P(r) over all r's, we can get an identity equation:

$$\Sigma_r (P(r)) = 1 \tag{2}$$

3. Parameter Estimation and Empirical Findings

3-1 POS Data of Beer Purchase

In order to estimate the parameter values, we make use of a set of POS data scanned at a convenience store. The store of a liquor shop type is located at Setagaya-ku, Tokyo, Japan. They have more than 500 card members who occasionally shop alcoholic drinks or beverages at the store.

Based on the customer's shopping frequency of more than 10 buying occasions for each customer, we extracted 139 members from the entire sample in a computer file. The purchase record had been kept on June 1 through December 2, 1991. Total number of purchase occasions is 1,890 during this time period. And the total number of units bought out by 139 members is 6,777 cans or bottles. Therefore, the average consumption rate of beer per member is approximately 49 cans (bottles) for five months.

Thirty-two brands were recorded at least once for any panel member. Core brand is to be defined as a brand which was most frequently bought by each customer. The panel shares of major five brands are tabulated in Table 2. Despite the frequent introduction of new brands into this market, concentration ratio of market shares is very high. Top five brands account for about 75% share in volume unit.

Table 2: Summary Statistics of Purchase Records

(1) Location of the panel CVS : Setagaya, Tokyo, Japan
(2) Recording period : June 1 to December 2, 1991

(3) The number of panelers available = 139
(4) Total purchase occasions = 1,890
(5) Total purchase units = 6,777

(6) Average purchase units per member = 48.8
(7) Average purchase occasions per member = 13.6
(8) Average purcahse units per occasion = 3.6
(9) Average purchase units per month = 9.8

(10) Market Share of Major five Brands

Brand Name	Purchase Units	Panel Share
Kirin's Lager	1,358	20.0 %
Asahi's Super Dry	1,222	18.0 %
Sappro Black Label	1,149	17.0 %
Ichiban-shibori	1,111	16.4 %
Asahi's Z	217	3.2 %
Other Brands	1,720	25.4 %
Total	6,777	100.0 %

3−2 Specification for Parametric Models

(1) Construction of Log−likelihood Function

Equations (1−n) to (1−0) have two parameters, i.e., π_A and α. Therefore, we can estimate the parameter values directly by maximizing log−likelihood function LL:

$$LL = \quad ((\ln D_n! - \quad \ln D_{rn}!) + \quad (Drn \ln P_n(r))) \qquad (3)$$

where,

$P_n(r)$: the chance probability in eq. (1) that a consumer choosesr units of his core brand out of total n trials at a purchase time,

D_n : the total number of purchase occasions when a consumer chooses n items at a purchase time,

D_{rn} : the number of purchase occasions when a consumer buys r units of his core brand at a time.

We maximize LL in eq. (3) over π_A and α to get an optimal set of parameters.

(2) A Logarithmic Transformation of Basic Parameters

Since we like to compare the parameter values with those calibrated in the other nested models, we slightly change the parameter specification as follows:

[Basic Model]

$$\pi_A = 1/(1+\exp(-b_0)) \qquad (4)$$

$$\alpha = 1/(1+\exp(-a_0)) \qquad (5)$$

where,

$0 < \pi_A, \alpha < 1.$

This logarithmic transformation gives us an advantage that parameters π_A and α lie only in the positive area. There is one-to-one correspondence between π_A and b_0 in eq. (4). This relationship is also true for the pair of α and a_0 in eq. (5). Notice that both choice probabilities, π_A and α, are constant over the number of units chosen.

(3) U-shape Hypothesis

Imagine that a consumer is going to buy beer at a convenience store. According to our assumptions in two step choice model, he has to decide whether he buys a single label in a bundle (Case I: bundle buying), or picks up several brands in a mixed match (Case II: mixed match buying).

In the former case I, α, which we call "bundle buying probability", must be relatively high with a low value of n (= 2 or 3). We can expect that α decreases with an increasing number of n, since there is a tendency towards variety seeking when a consumer has an opportunity to buy more than a couple of cans of beer (n = 3, 4, or 5).

Interestingly, however, the bundle buying probability α may increase again when the number of buying units gets closer to six (n = 6). This resurgence of α value may be caused partly by the promotional efforts of beer brewing companies. They sell beer in a six pack or in a carton box (n = 20). Another reason is that heavy buyers would be much more brand loyal than the middle or light buyers.

In order to test this U-shape hypothesis for α, we specify a quadratic function of n for the bundle purchase parameter α :

[U-shape Hypothesis]

$$\pi_A = 1/(1+\exp(-b_0)) \qquad (4)$$

$$\alpha = 1/(1+\exp(-(a_0+a_1 n+a_2 n^2))) \qquad (5)'$$

where,

$0 < \pi_A, \alpha < 1.$

3-3 Parameter Estimation

Using the POS data described above, we obtained a set of parameter estimates for both basic model and U-shape hypothesis model. Maximum-likelihood estimation procedure was used to derive parameter values for three different kinds of data sets: (1) the entire population, (2) three sub-samples on loyalty base (High loyalty group, Middle loyalty group, Low loyalty group), and (3) five sub-samples on core brand base (Kirin's "Lager", Asahi's "Dry", Kirin's "Ichiban-Shibori", Sapporo's "Black-Label", Asahi's "Z").

(1) Parameter Estimation for Basic Model

In the first column of Table 3, we can see the "average" zero-order choice probability π_A and "flat" bundle purchase probability α . In actual, we did not use the whole samples, but discarded the purchase data if $n > 6$. Unfortunately, probability prediction gets very unstable for $n > 6$ because of few purchase incidence over this range.

Estimated α value is at around .7, which means that an average consumer is expected to buy a single sort of beer brand seven times out of ten purchase opportunities. In turn, a consumer seeks for a variety with an odd of 30 percent. The zero-order probability π_A was estimated to be a little bit more than half. This figure .535 might be surprising, since people believe that beer drinkers tend to switch from one to another brand quite often.

Table 3: Parameter Estimates for All Panelers

Parameter	Basic Model	U-Shape Model
α	.699	1.000 (r=1)
		.720 (r=2)
		.663 (r=3)
		.653 (r=4)
		.692 (r=5)
		.770 (r=6)
πA (=P1)	.535	.535
b0	.142	.142
a0	.845	2.136
a1		-.818
a2		.116
ρ2	.455	.473
χ2	-	7.969

Substituting these values into eq. (1), we can predict the choice probability of core brand as a function of chosen unit r. The choice probabilities for the theoretical distribution are shown in Table 4 for n = 2 to 6. A comparison can be made between the actual and estimated choice probability distributions for each number of unit chosen at a time (Table 4).

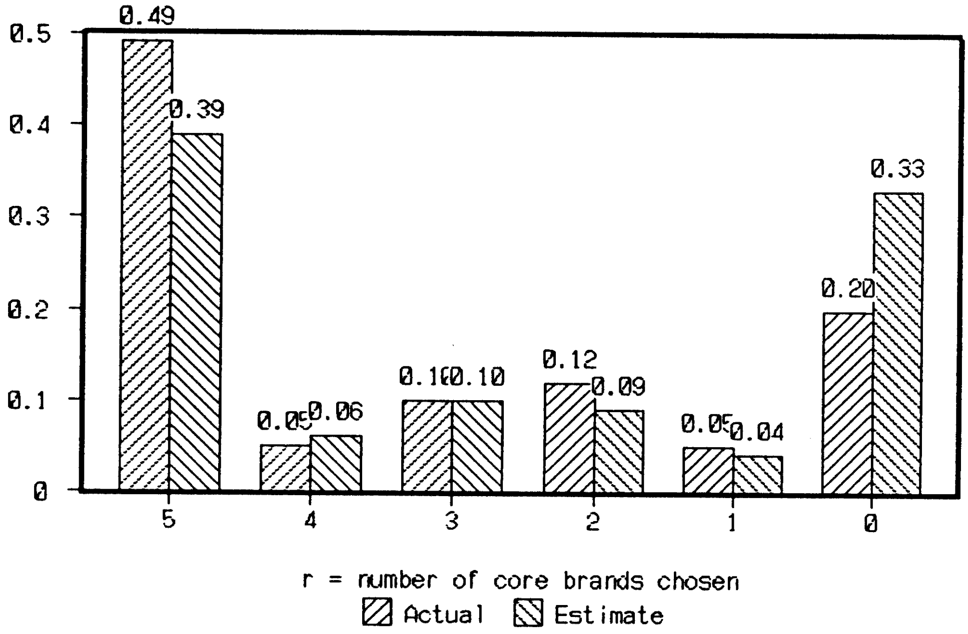
For an illustrative example in Figure 3, a graphical presentation is given for n = 5. Internal validity of our choice model looks like very nice. However, there are two exceptions at both ends (r = 0 and 5) of W– shape distribution.

Table 4: Probability Fitting

n = 2	r =	2	1	0				
Actual		0.52	0.14	0.34				
Estimate	Basic Model	0.46	0.15	0.39				
	U-Shape	0.47	0.14	0.40				
	Adjusted	0.52	0.14	0.34				
n = 3	r =	3	2	1	0			
Actual		0.43	0.12	0.14	0.31			
Estimate	Basic Model	0.42	0.12	0.10	0.36			
	U-Shape	0.41	0.13	0.12	0.34			
	Adjusted	0.44	0.13	0.12	0.31			
n = 4	r =	4	3	2	1	0		
Actual		0.48	0.04	0.19	0.04	0.24		
Estimate	Basic Model	0.40	0.09	0.11	0.06	0.34		
	U-Shape	0.38	0.10	0.13	0.07	0.32		
	Adjusted	0.47	0.10	0.13	0.07	0.23		
n = 5	r =	5	4	3	2	1	0	
Actual		0.49	0.05	0.10	0.12	0.05	0.20	
Estimate	Basic Model	0.39	0.06	0.10	0.09	0.04	0.33	
	U-Shape	0.38	0.06	0.10	0.09	0.04	0.33	
	Adjusted	0.51	0.06	0.10	0.09	0.04	0.20	
n = 6	r =	6	5	4	3	2	1	0
Actual		0.67	0.01	0.09	0.06	0.06	0.01	0.12
Estimate	Basic Model	0.38	0.04	0.08	0.09	0.06	0.02	0.33
	U-Shape	0.42	0.03	0.06	0.07	0.05	0.02	0.36
	Adjusted	0.66	0.03	0.06	0.07	0.05	0.02	0.12

Remarks: * n = number of units purchased
** r = number of core brands chosen

Figure 3: Actual v.s. Prediction Probs: Basic Model (n = 5)



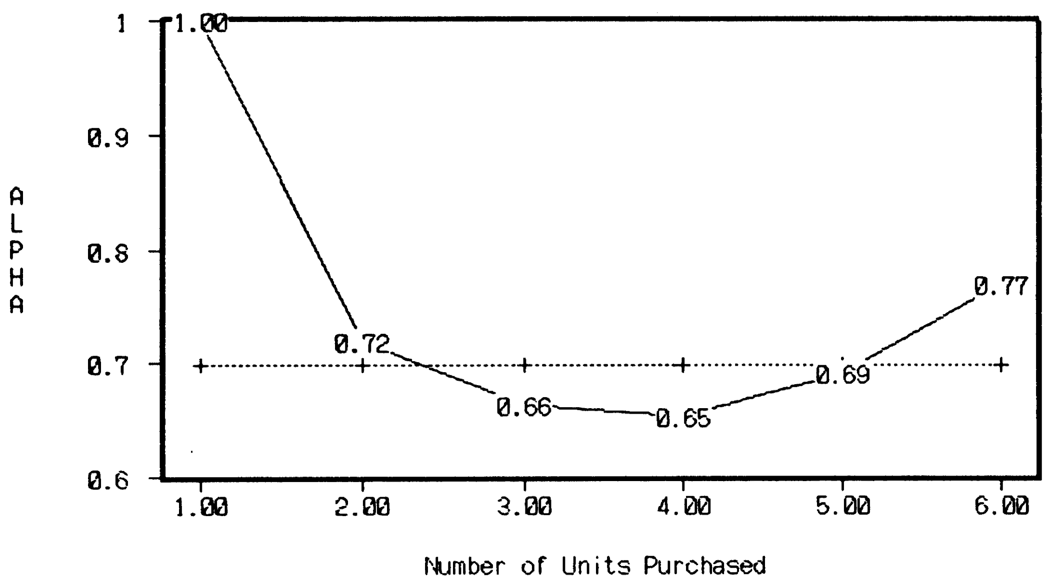
(2) Parameter Estimation for U-shape Hypothesis Model

Two additional parameters, a_1 and a_2, were estimated for the U-shape hypothesis. We can find these new parameters in the second column of Table 3.

As was expected, "aggregate" α function as of an argument n (the number of units chosen) resulted in a U-Shape curve with $a_0 = 2.14$, $a_1 = -.82$, $a_2 = .11$. Note that parameters a_0 and b_0 for Basic Model are .85 and .14, respectively. Both bundle purchase probability curves are overlaid in Figure 4 for comparison.

This U-shape pattern is not necessarily true for the sub-samples. Rather, as it will be seen in the following sub-section, we will figure out a reversed U-shape curve for brands with less loyalty.

Figure 4: U-shaped ALPHA function For All Panelers



Fitting measure of our brand choice model can be calculated by using McFadden's ρ^2 : (*9)

$$\rho^2 = 1 - LL(a_0, b_0 = 0) / LL(a_0, a_1, a_2, b_0) \qquad (6)$$

where,

LL (.) corresponds to an optimized log-likelihood given estimated parameters.

Also in comparing the nested models, Basic and U-shape Models, an additional contribution of parameters to fitting measure by adopting U-shape Model can be evaluated by using Rao's χ^2 -square measure:

$$\chi^2 = 2\,(LL\,(a_0, a_1, a_2, b_0) - LL\,(a_0, b_0 = 0)) \qquad (7)$$

In case of Basic Model, ρ^2 is .455, which could be improved to be .473 by introducing a variable α function of n. The corresponding χ^2 is 7.97 with two degrees of freedom. Since the critical value of χ^2 -square distribution is 6.0 with p = .05 for d.f. = 2, we can insist that the improvement by adopting the U-shape function is significant at five percent level.

Minimum value of α is attained as .65 .at r = 4. The maximum value is .77 at r = 6. The difference between the minimum and maximum values of bundle purchase probability is to be about .12. It is said that there is certainly a bundle buying effect.

3-4 Empirical Findings for Sub-Groups

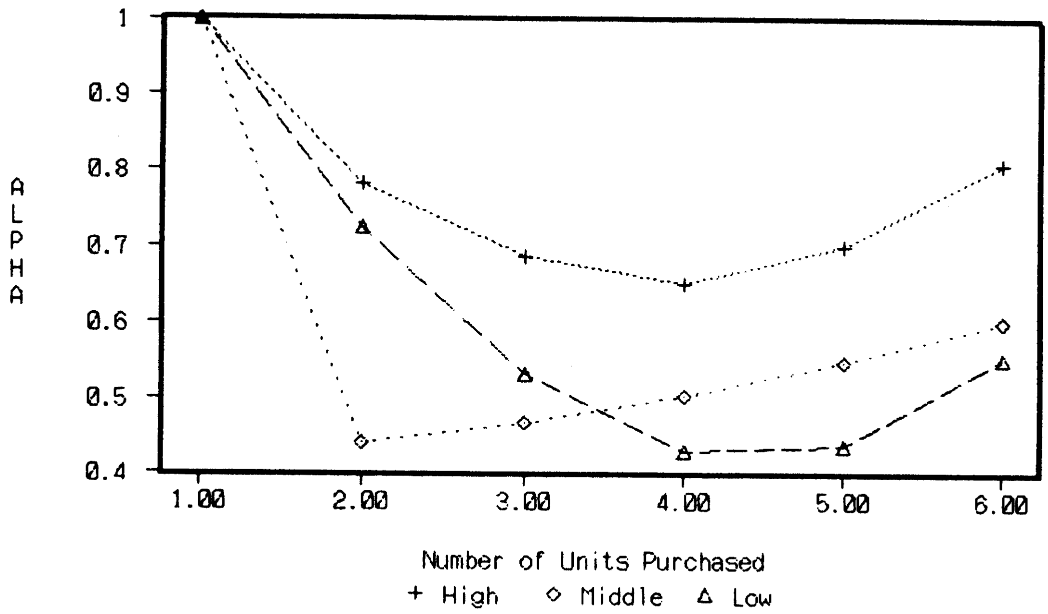
In this sub-section, we split the entire purchase data into sub-samples. This file splitting is performed on consumer base in two ways: (A) by purchase loyalty (for High, Middle, Low loyalty groups), and (B) by core brand (for five major brands: Kirin's "Lager", Asashi's "Dry", Kirin's "Ichiban-Shibori", Sapporo's "Black-Label", Asahi's "Z").

After calibrating the estimate values by ML method, we examined the difference in multiple brand purchase behavior among consumer segments. Estimated parameters of π_A and α by segment enabled us to interpret the special features of bundle purchase pattern by sub-samples.

(1) Bundle Purchase Probability: α

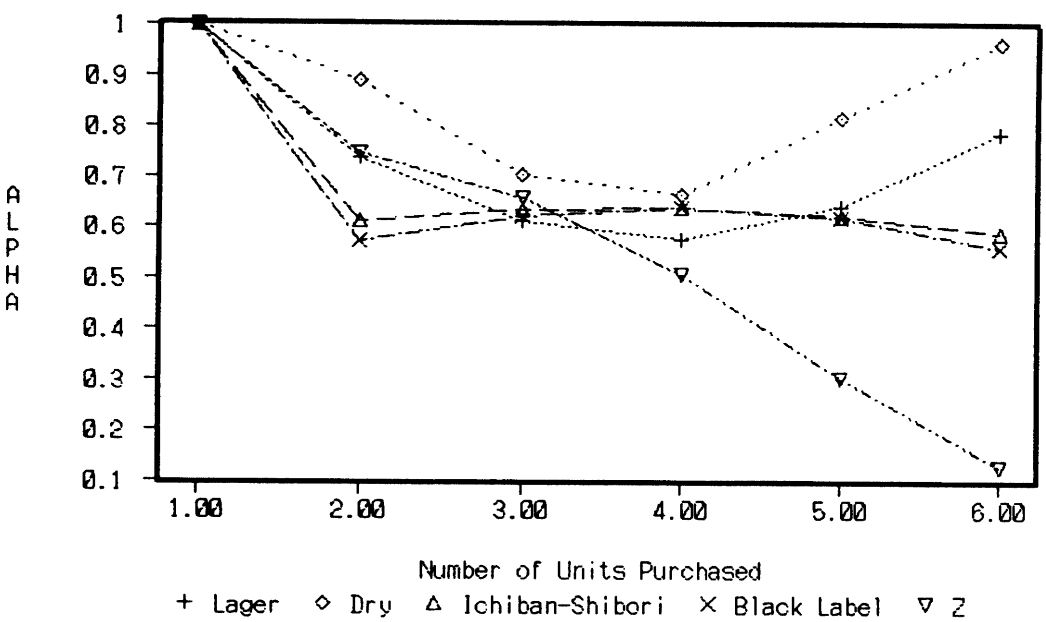
Estimated α parameters are shown in Figure 5 along with an axis of the number of units purchased at one time. Of interest is that the bundle purchase probability α shows a very similar pattern between the high and low loyalty groups, although its absolute degrees of purchase loyalty are quite different. Actually, a consumer group with the higher loyalty has much higher probability than that with low loyalty almost by sixty percent.

Figure 5: U-shaped ALPHA function: By Loyalty



In contrast, α value in the middle loyalty group is monotonically increasing with an increasing number of purchase units. A consumer who has a moderate loyalty to his core brand tends to make a single brand purchase less often at low level of purchase unit (n = 2, 3). It means that they like to seek for a variety with a few cans of beer on a multiple items selection occasion.

Figure 6: U-shaped ALPHA function: By Brand



Next we estimate the α parameters by core brand. In Figure 6, we can identify three distinctive clusters of α functions, depending on the signs of a_1 and a_2, i.e., the shape of bundle purchase probability curve.

Two major brands with high market share, Kirin's "Lager" and Asahi's Dry, have a U-shape function of n. The Kirin's "Ichiban-Shibori" and Sapporo's "Black Label" with a moderate panel share constructs another cluster with a reverse U-shape, which looks like a flat hill. The lowest share brand, Asahi's Z, shows a upside-down J-shape. It seems to us that consumers relatively loyal to "Z" may not buy it more often when they want to drink more beer.

This pattern analysis suggests that the number four of purchase unit looks like a magic number. With an exception of Asahi's "Z", all other curves of bundle purchase probability have a maximum or minimum at the number of unit four.

An alternative explanation is that, if the brand lacks of loyalty, then there is less opportunity for it to be promoted at a store in a bundle. In contrast, the high share brand can sell well because it gets a lion's share in a bundle sales.

(2) Choice Probability of Core Brand: π_A

Also we can estimate the choice probability π_A on sub-sample base, as well as for the entire population. In Table 5 we show a set of parameter estimates for eight segments: three for loyalty segment, five for core brand segment. Not only π_A but also α and panel shares are presented in that table for the reference purpose.

Readers may be interested in the section by core brand. It is amazing that the value of π_A is not necessarily proportional to that of α or panel share. Zero-order choice probability of a core brand with the higher share might be less than that with the lower share. This is the case for "Lager" to ""Dry, as well as for "Ichiban-Shibori" to "Black Label".

To building up its market share, any brand has a major source from which it can obtain most customers or sales. Table 5 suggests for the possibility that, if consumers purchase multiple items at a time, we could find out some counter-examples for double jeopardy (Ehrenberg and Goodhardt 1990) in a different context. (*10) That is, high share brand may not enjoy the high purchase rate, which corresponds to α or π_A in our case.

Table 5: Parameter Estimates by Segment for Basic Model

Segment	α	π_A	Share
All Panelers	.699	.535	
By Loyalty			
High	.734	.879	
Middle	.508	.518	
Low	.564	.201	
By Brand			
Lagar	.674	.669	.200
Dry	.818	.755	.180
Ichiban-S	.619	.434	.170
Black Label	.600	.559	.164
Z	.502	.410	.032

(3) Adjustment of purchase probability: P_1

As was suggested in the preceding sub-section, theoretical choice probability of core brand for the number of units chosen, $P(r|n)$ in eq. (1), tends to be underestimated for $r = 0$ and n. It is said that zero-order purchase probability π_A might have a premium at both ends ($r = 0$ and n), compared to that in case II, i.e., mixed purchase occasion in the inner region of r. In other words, zero-order probability π_A has to be adjusted to become larger when a consumer purchases beers of a kind in a bundle.

Thus, we put an additional parameter g_0 to eq. (4), which results in the following equation:

[Adjusted π_A Model]

$$\pi_A = 1/(1+\exp(-(b_0+g_0))) \tag{4)'}$$
$$\alpha = 1/(1+\exp(-(a_0+a_1 n+a_2 n^2))) \tag{5)'}$$

where,

$g_0 = 1$ if CASE I, (then $\pi_A = P_1$)

$\quad 0$ if CASE II, (then $\pi_A = P_0$),

$0 < \pi_A, \alpha < 1$.

With an above correction, we can reevaluate the projected probability of core brand. The bar chart shows a much better fit of theoretical probability in Figure 7 (for $n = 5$) after an adjustment of π_A in bundle buying. For the other cases of $n = 1$ to 6, the fit could be also improved. We have McFadden's ρ^2 of .702.

In Figure 8, we show a zero-order purchase probability curve π_A along with the number of units chosen. It is hard to say there is a clearly identified among the three lines as was in the α curve. Likewise, we cannot insist of any general statements for Figure 8, in this case by core brand.

One can guess that the more units consumers buy, the more zero-order choice probability of core brand they might have. However, there are a couple of exceptions in a mixed bundle of items, e.g., for "Ichiban-Shibori" and "Z" at the number of unit chosen of six.

Figures 7: Actual v.s. Prediction Probs: Adjusted

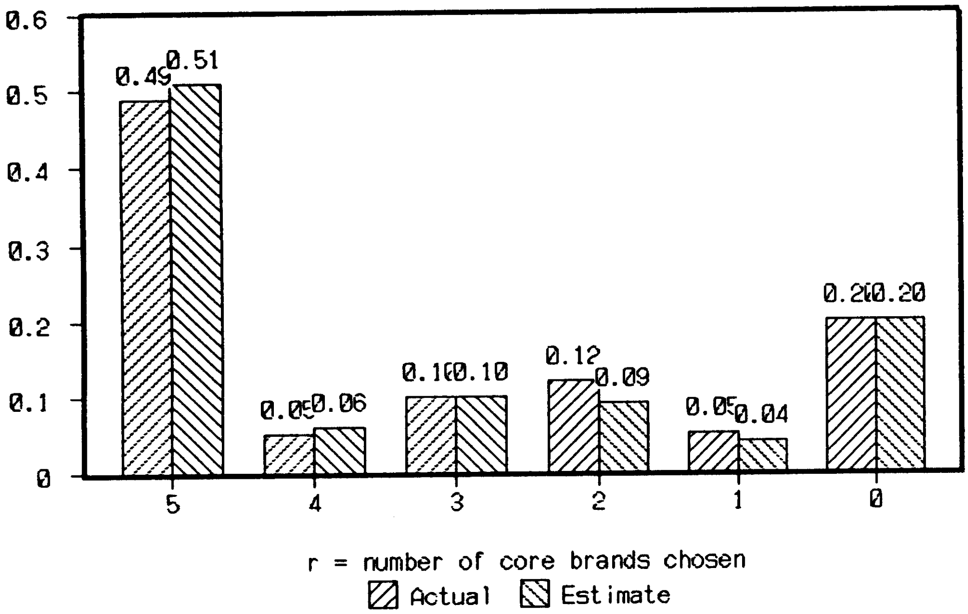


Figure 8: Estimated PA: By Loyalty

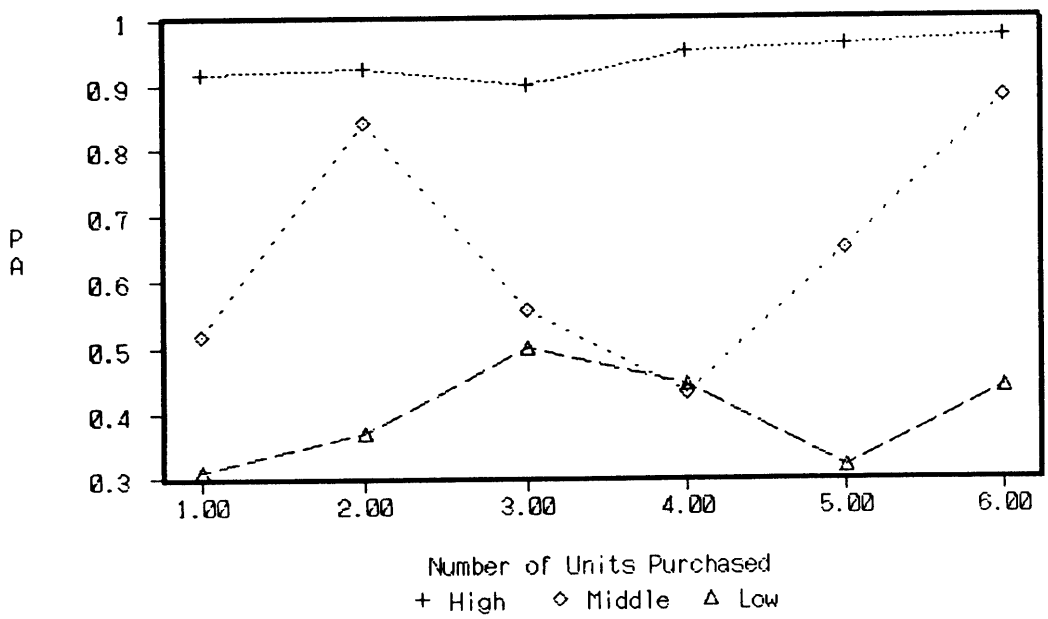
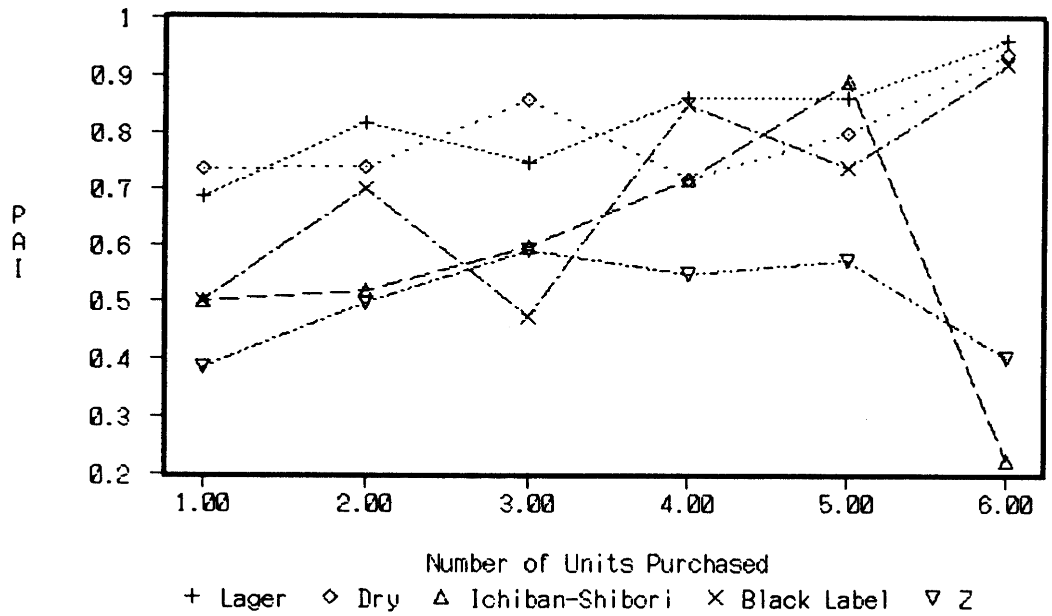


Figure 9: Estimated PA: By Brand



4. Discussion and Conclusions

We summarize our analysis as follows:

(1) We identified a unique phenomenon of multiple items selection behavior in some product categories, and successfully tried to construct a two-stage model of brand choice.

(2) In the model, choice probability of core brand (the most favorite) could be decomposed into two elements: one for bundle purchase of a single brand (with probability of α), and the other of mixed buying of several items. We assumed a zero-order purchase probability π_A in choosing the number of core brands.

(3) Using POS data of beer purchase record scanned at a convenience store, we estimated π_A and α parameters by ML method for all panelers. Splitting the entire samples into sub-samples, one by loyalty intensity and the other by core brand, we compared the parameter values across consumer segments.

(4) A choice model with variable bundle purchase probability led us to a couple of interesting results. One of them is the aggregate U-shape curve. We also obtained a magic number four, at which the choice probability of core brand tend to be optimal in a multiple choice situation.

(5) With calibrated estimates by brand or segment, we have an alternative way to understand

the relative market position of brands in a competitive market.

Finally we like to point out four frontiers, which we did not deal with in this paper, but certainly have an opportunity to extend for a research study:

(1) To extend our simultaneous items selection model into the more general frame work, i.e., in case of more than two brands.

(2) To include consumer's attributes in our choice model, which may lead to a logit-type formulation of great popularity in an analysis of scanner data.

(3) Similar to the second one, to include marketing mix variables: price, promotion, GRP, and so forth, in data analysis.

(4) To be able to specify a dynamic feedback model for consumer's brand choice because ours is a sort of variety-seeking model. (*11)

NOTES

(*1) Ehrenberg, A.S.C. (1988), *Repeat Buying, 2nd ed.*, New York: Oxford University Press. Also we can see a typical orientation by British scholars in the work by G.J. Goodhardt and A.S.C. Ehrenberg (1967), "Conditional Trend Analysis: A Breakdown by Initial Purchase Level," *Journal of Marketing Research, 4 (May),* 155-161.

(*2) Very often Professor Donald Morrison co-authored with his students. Such joint works are as follows: Schmittlein, D. C., D.G. Morrison and R.A. Colombo (1987) "Counting Your Customers: Who are They and What will Do Next?" Management Science, 33 (January), 1-24. Colombo R.A. and D.G. Morrison (1989), "A Brand Switching Model with Implications for Marketing Strategies," *Marketing Sciences, 8 (Winter),* 89-99.

(*3) First application of Logit model to POS data appeared in Marketing Science by Guadagni P.M. and J.D.C. Little (1988),"A Logit Model of Brand Choice Calibrated on Scanner Data,"*Marketing Science, 2 (Summer),* 203-238.

(*4) A comprehensive review of variety-seeking behavior in choice modeling up to 1980s was written by McAlister and Pessimier in 1982: McAlister L. and Pessemier (1982), "Variety Seeking Behavior: An Interdisciplinary Review," *Journal of Consumer Research, 9 (December),* 311-322. More general definition of variety-seeking behavior was done by Kahn, Kalwani, Morrison in 1986. Kahn B. E., M. U. Kalwani, and D.G. Morrison (1986), "Measuring Variety-Seeking and Reinforcement Behavior Using Panel Data," *Journal of Marketing Research, 23 (May),* 89-100.

(*5) Massy W.F., D.B. Montgomery, and D.G. Morrison (1970), *Stochastic Models of Buying Behavior, MIT Press (Part II).*

(*6) Nakajima, N. (1993), "Measuring Price Elasticities Using Scanner Data," in Ogawa K. ed., *POS and Marketing Strategy, Tokyo: Yuhikaku,* 257–280.

(*7) Ignoring the first step, we call this model as a two step approach later on.

(*8) Gupta, S. (1991), "Stochastic Models of Interpurchase Time with Time–Dependent Covariates," *Journal of Marketing Research, 28 (February),* 1–15.

(*9) McFadden, D (1980), "Econometric Models for Probabilistic Choice Among Products," *Journal of Business, 53, Part 2 (July),* 513–530.

(*10) Ehrenberg, A.S.C. et al (1990), "Double Jeopardy Revisited," *Journal of Marketing.*

(*11) A dynamic specification of variety seeking behavior can be seen in a paper by Lattin and McAlister. Lattin J.M. and L. McAlister (1985), "Using a Variety–Seeking Model to identify Substitute and Complementary Relationships Among Competing Products," *Journal of Marketing Research, 22 (August),* 330–339.