数理解析研究所講究録
第1804巻 2012年 144-161

The Latent Class Model of Brand Choice Behaviors Incorporating Variety-Seeking and State Dependence.

April 24, 2012

Graduate School of Systems and Information, University of Tsukuba.
Tomohito Kamai
Department of School of Systems and Management, University of Tsukuba.
Yuichiro Kanazawa.

1 Introduction

Understanding the dynamic consumers' brand choice behaviors is an interest to both retailers and manufacturers who plan their marketing strategies. One of the major topic in this field includes the behavioral patterns seemingly persisting across more than one purchase occasion. They are referred to as variety-seeking and state dependence.

The term "variety-seeking" is defined as the behavior such that the purchase of a brand will decrease the probability that the same brand will be purchased again on the succeeding occasions. The "state dependence", sometimes called "inertia", refers to the opposite behavior; the brand purchase will increase the probability of purchasing the same brand again on the succeeding occasion (Bawa, 1990; Chintagunta, 1998).[1]

The existence of variety-seeking and state dependence plays an important roles in marketer's decision making as Chintagunta (1998) states: "From a managerial stand point, it would be important to know whether a brand's consumers are inertial or variety prone (Chintagunta, 1998, 254)." For example, the existence of a strong state dependence effect suggests that inducing trial would be an effective marketing tactic; the promotional schemes such as a free sampling would be more effective. The emphasis on brand retention would be also important to marketers of this type of products (Chintagunta, 1998). On the other hand, the existence of variety-seeking behavior motivates marketers to extend their product lines so that households' brand switching behaviors would benefit their own products (Seetharaman, 2004).

In summary, by correct understanding of the consumers' brand choice behaviors, the promotions can be effectively implemented to the right consumers with the right schemes. In this study, we will propose a model to capture the complex behaviors of consumers who vary their choices without any apparent reasons.

2 Literature Review

According to the taxonomy work on literature of varied behaviors of McAlister and Pessemier (1982), the variety-seeking behavior may not necessarily be associated with a motivation for variety-seeking itself; this behavior is often observed, for instance, when a household has multiple users for the same product. The "real" or "pure" variety-seeking behavior we consider in this study is a varied behavior motivated by a desire for an unfamiliar product or a stimulus associated with switching behavior. In this context, Givon (1984)

---

[1] When the purchase probability is affected by the last purchase (i.e., Markov process), it is called "first-order" behavior. Accordingly, "zero-order" behavior refers to the pattern in which the purchase probability is not affected by any previous choices.

used a stochastic model to express a utility associated with the switching behavior itself in addition to the utility derived from the consumption of the specific brands.

In contrast to variety-seeking behavior, some researchers found the "inertial behavior", which is often referred to as "state dependence." One of the most influential paper in this stream is Guadagni and Little (1983), who specified variables representing brand and size loyalties in the way that authors assumed weighted sequential influences of past choices to the current utility in the form of

$$\begin{aligned} x_{ijt_i} &= \alpha_b \cdot x_{ij,t_i-1} \\ &\qquad +(1-\alpha_b)\{\textit{consumer } i \textit{ bought brand } j \textit{ at } (t_i-1)\textit{-th occasion}\}, \end{aligned}$$

where α_b is a parameter and $\{\textit{statement}\}$ denotes an indicator function taking the value unity if the statement is true.

Bawa (1990) proposed a "hybrid" behavior, which is characterized by the behavior where consumers seem to exhibit inertial tendency for certain period of time and then exhibit variety-seeking tendency once certain period of time passes. The marginal utility in his model is specified to be the function of "run", denoted by $r_{ij}(t_i)$, which is the number of times the brand had been continuously purchased up to t_i-th occasion. The index i and j denote consumer and product respectively.

3 The Specification of the Utility and the Model

In constructing the model, we choose to use the brand loyalty variable of Guadagni and Little (1983), which we will refer to it as "GL variable" henceforth, to express the state dependence part of the utility. Also, to capture the effect of variety-seeking tendency, we choose to include run which was defined in Bawa (1990). The purpose of including run is to "put brake" on the GL variable, which keeps increasing as long as the same brand is kept purchased. By including run, the consumer's utility for the brand would start to decline as a result of the repeated consumption of the same brand if run negatively affects utility.

Now if consumer is state dependent, the effects of both GL variable and run will be non-negative. The variety-seeking behavior will be detected by the opposite signs; the effects of both GL variable and run will be non-positive. On the other hand, the hybrid behavior will be detected by the positive coefficient of GL variable and negative coefficient of run. Since the combination of positive coefficient of GL variable and negative coefficient of run could indicate both variety-seeking and hybrid behaviors, depending on the values of coefficient, we must scrutinize the results carefully before we judge.

Now we write the utility of consumer $i = 1, \ldots, N$ for brand $j = 1, \ldots, J$ at occasion $t_i = 1, \ldots, T_i$ as

$$U_{ijt_i} \quad = \quad \boldsymbol{x}_{ijt_i}\boldsymbol{\beta}_s + \epsilon_{ijt_i}, \tag{3.1}$$

where $\boldsymbol{x}_{ijt_i}$ is $1 \times R$ vector of the explanatory variables including a set of dummy variables for brands except for one base brand, dummy variables for coupon usage, the average coupon values redeemed, dummy variables for feature and display, GL variable and run. The $\boldsymbol{\beta}_s$ is corresponding $R \times 1$ vector of parameters for segment s. The segment is a subset into which consumers are placed, where consumers in the same segment are assumed to have homogeneous characteristics regarding preferences for brands and responsiveness to the marketing variables. In our framework, we also assume that consumers in the same segment show the same purchasing patterns expressed in GL variable and run. The term ϵ_{ijt_i} is a random error term that captures the effects of unobserved variables which follows i.i.d. Gumbel distribution.

In our study, we employ the latent class model which is one of the general models to incorporate heterogeneity across consumers. In the latent class model, it is assumed that there exist a finite and fixed number of segments where each consumer belongs to only one of the segments, and it is further assumed that consumers belong to the same segment over the period of observation. The idea behind this type of model is that there exists an underlying multi-dimensional distribution of parameters for intrinsic preferences for brands and relative responsiveness to the marketing variables which characterize consumers' behaviors. In the latent class model, the underlying distribution of parameter is assumed to be discrete. Because the finite representation of consumers' characteristics of the latent class model coincides well with the concept of segments, such a model is widely applied to the marketing field. The major work of this field is Kamakura and Russel (1989). The other works using the latent class model include Bucklin et al. (1998) and Gupta and Chintagunta (1994).

Now we will use the multinomial logit model framework for brand choice, and the log likelihood of panel data in our model is given by

$$l(\boldsymbol{\beta}) = \sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{t_i=1}^{T_i}\sum_{j=1}^{J}\left\{h_i(s)\cdot y_{ijt_i}\cdot \ln\left(\frac{\exp(\boldsymbol{x}_{ijt_i}\boldsymbol{\beta}_s)}{\sum_{l=1}^{J}\exp(\boldsymbol{x}_{ilt_i}\boldsymbol{\beta}_s)}\right)\right\}$$

where y_{ijt_i} is the indicator variable taking value 1 if consumer i buy brand j at occasion t_i and $h_i(s)$ is the expected value of membership probability of consumer i to segment $s = 1, \ldots, S$. For estimation, we used EM algorithm assuming the consumers' membership probabilities for each segment as missing values.

4 Estimation

Now let us assume there are $s = 1, \ldots, S$ segments in the sample. Obviously as the number of segments is unknown nor can be observed, it must be estimated. First define the relative sizes of segment s as π_s such that

$$0 < \pi_s \leq 1$$

for all s and

$$\sum_{s=1}^{S} \pi_s = 1. \tag{4.1}$$

In the latent class model, each consumer has different membership probabilities for these segments, because membership probabilities are estimated from their choice histories which differ across consumers. Now let us define the random variable Y_{ijt_i} for $i = 1, \ldots, N$ and $t_i = 1, \ldots, T_i$, which takes value one if consumer i chooses brand j at t_i-th occasion. In other words, for consumer i, let y_{ijt_i} be entries of $T_i \times J$ matrix $\boldsymbol{Y}_i$

$$\boldsymbol{Y}_i = \begin{pmatrix} y_{i11}, & \cdots & , y_{iJ1} \\ \vdots & \vdots & \vdots \\ y_{i1T_i}, & \cdots & , y_{iJT_i} \end{pmatrix} \tag{4.2}$$

and let us denote each row as $\boldsymbol{y}_{it_i}$. Assuming the ϵ_{ijt_i} follows i.i.d. extreme value with respect to j, we can express the probability that consumer i in segment s chooses brand j at the t_i-th occasion in the standard logit form as

$$\Pr\{(y_{i1t_i}, \ldots, y_{iJt_i}) = (\underbrace{0, \ldots, 0}_{j-1}, 1, \underbrace{0, \ldots, 0}_{J-j}) | S_i = s; \boldsymbol{\beta}_s\} = \frac{\exp(\boldsymbol{x}_{ijt_i}\boldsymbol{\beta}_s)}{\sum_{l=1}^{J} \exp(\boldsymbol{x}_{ilt_i}\boldsymbol{\beta}_s)}, \tag{4.3}$$

where the random variable S_i indicates which segments consumer i belongs to, assuming we could observe the segment membership of consumer i. For brerity, we abbreviate (4.3) as

$$\Pr(Y_{it_i} = j | S_i = s; \boldsymbol{\beta}_s) = \frac{\exp(\boldsymbol{x}_{ijt_i}\boldsymbol{\beta}_s)}{\sum_{l=1}^{J} \exp(\boldsymbol{x}_{ilt_i}\boldsymbol{\beta}_s)}, \tag{4.4}$$

henceforth.

The unconditional choice probability for brand j of a randomly selected consumer i can be obtained by integrating out the equation (4.3) by the density in the population π_s as[2]

$$\Pr(Y_{it_i} = j) = \int \Pr(Y_{it_i} = j | S_i = s; \boldsymbol{\beta}_s) \cdot \pi_s ds. \tag{4.5}$$

Since the relative size of the segment π_s is discrete, (4.5) is written as

$$\Pr(Y_{it_i} = j) = \sum_{s=1}^{S} \pi_s \cdot \Pr(Y_{it_i} = j | S_i = s; \boldsymbol{\beta}_s). \tag{4.6}$$

This is a weighted average of logit formula evaluated at each mass point (segment), as pointed out by Kamakura and Russell (1989).

Now suppose that consumer i has the choice history defined as $H_i = (H_{i1}, \ldots, H_{iT_i})$, where H_{it_i} indicates the brand purchased at t_i-th occasion. Then the conditional choice probability that consumer i has the choice history H_i given that consumer i belongs to segment s is written as

$$\Pr(H_i | S_i = s; \boldsymbol{\beta}_s) = \prod_{t_i=1}^{T_i} \prod_{j=1}^{J} \{\Pr(Y_{it_i} = j | S_i = s; \boldsymbol{\beta}_s)\}^{y_{ijt_i}}. \tag{4.7}$$

In the same manner as (4.6), the unconditional probability of randomly selected consumer i having the choice history H_i can be written as

$$\Pr(H_i; \boldsymbol{\beta}) = \sum_{s=1}^{S} \pi_s \cdot \Pr(H_i | S_i = s; \boldsymbol{\beta}_s) \tag{4.8}$$

where $\boldsymbol{\beta}$ is $R \times S$ parameter matrix

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_S) = \begin{pmatrix} \beta_{11}, \ldots, \beta_{1s}, \ldots, \beta_{1S} \\ \vdots \\ \beta_{r1}, \ldots, \beta_{rs}, \ldots, \beta_{rS} \\ \vdots \\ \beta_{R1}, \ldots, \beta_{Rs}, \ldots, \beta_{RS} \end{pmatrix}. \tag{4.9}$$

Now if the segment memberships of consumers are completely known, the vector of parameters $\boldsymbol{\beta}_s$ can be estimated by the well known methods such

[2]The model of the form (4.5) is sometimes called mixed logit model and π_s is called mixing distribution. The latent class model can be regarded as the special case of mixed logit model where mixing distribution is discrete (Train, 2003).

as Newton-Raphson method. Here let us define for each i the multinomial indicator random variable $z_i(s)$ which takes one if consumer i belongs to segment s and 0 otherwise, assuming we know the membership probability of consumer i belonging to segment s given his/her purchase history H_i, that is $\Pr(S_i = s|H_i; \boldsymbol{\beta}_s)$. Then this membership indicator random variables $z_i(s)$'s are entries of $N \times S$ matrix $\boldsymbol{Z}$ as

$$\boldsymbol{Z} = \begin{pmatrix} z_1(\cdot) \\ \vdots \\ z_N(\cdot) \end{pmatrix} = \begin{pmatrix} z_1(1), \ldots, z_1(S) \\ \vdots \\ z_N(1), \ldots, z_N(S) \end{pmatrix}.$$

The row sums of the matrix $\boldsymbol{Z}$ above are all 1.

Assuming we were able to observe $\boldsymbol{Z}$, the likelihood given the choice histories of the all consumers under consideration is written as[3]

$$L(\boldsymbol{\pi}, \boldsymbol{\beta}|\boldsymbol{H}, \boldsymbol{Z}) = \prod_{i=1}^{N}\prod_{s=1}^{S} \{\pi_s \cdot \Pr(H_i|S_i = s; \boldsymbol{\beta}_s)\}^{z_i(s)},$$

where $\boldsymbol{H} = (H_1, \ldots, H_i, \ldots, H_N)$ is the choice history of all consumers in the sample, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_S)$ is $1 \times S$ vector of relative sizes of segments. Accordingly, the log likelihood could be written as

$$\begin{aligned} l(\boldsymbol{\pi}, \boldsymbol{\beta}|\boldsymbol{H}, \boldsymbol{Z}) &= \sum_{i=1}^{N}\sum_{s=1}^{S} z_i(s) \cdot \ln\left(\pi_s \cdot \Pr(H_i|S_i = s; \boldsymbol{\beta}_s)\right) \\ &= \sum_{i=1}^{N}\sum_{s=1}^{S} z_i(s) \cdot \ln \Pr(H_i|S_i = s; \boldsymbol{\beta}_s) + \sum_{i=1}^{N}\sum_{s=1}^{S} z_i(s) \cdot \ln \pi_s. \end{aligned} \tag{4.10}$$

Now if we were able to observe $\boldsymbol{Z}$, we can estimate parameters by the traditional method. However, in reality, we cannot obtain the information $z_i(s)$. In such a situation, the method called EM algorithm may be implemented to obtain the estimate of $z_i(s)$ along with the estimates of $\boldsymbol{\pi}$ and $\boldsymbol{\beta}$ as explained in the following subsection.

[3]The term $\pi_s \cdot \Pr(H_i|S_i = s; \boldsymbol{\beta}_s)$ is the joint probability that consumer i belongs to segment s and has choice history H_i. Note, however, that the relative size of segment π_s is unknown and has to be estimated.

4.1 EM algorithm

If the segment membership of consumers $\mathbf{Z}$ were completely known, the $\boldsymbol{\beta}_s$ can be estimated by the algorithm described above. EM algorithm takes advantage of this fact and in the algorithm, the consumer's membership to the segment $z_i(s)$ is assumed at first to be missing values and this value is imputed by its "expectation" (to be explained below). Then the conditional likelihood is maximized based on the expected value of membership to the segment. The consumers' expected membership is then updated using the updated likelihood. This cycle of "expectation" of membership to the segment and "maximization" of likelihood is repeated until the likelihood converges.

Now taking the expectation with respect to $z_i(s)$ for the log likelihood (4.10), we have

$$E[l(\boldsymbol{\pi},\boldsymbol{\beta}|\boldsymbol{H},\boldsymbol{Z})] = \sum_{i=1}^{N}\sum_{s=1}^{S} h_i(s)\cdot \ln \Pr(H_i|S_i=s;\boldsymbol{\beta}_s) + \sum_{i=1}^{N}\sum_{s=1}^{S} h_i(s)\cdot \ln \pi_s, \tag{4.11}$$

where

$$h_i(s) = E[z_i(s)] = \sum_{l=1}^{S} z_i(l)\cdot \Pr(S_i=l|H_i;\boldsymbol{\beta}_l) = \Pr(S_i=s|H_i;\boldsymbol{\beta}_s) \tag{4.12}$$

is the expected values of the indicator random variable $z_i(s)$ for $s=1,\dots,S$. Since parameter $\boldsymbol{\beta}$ in (4.9) only appears in the first term and $\boldsymbol{\pi}$ only appears in the second term on the right hand side of (4.11), they can be estimated by maximizing $E[l(\boldsymbol{\pi},\boldsymbol{\beta}|\boldsymbol{H},\boldsymbol{Z})]$ alternately.

Let us first look at the second term on the right hand side of (4.11). Since we have the condition $\sum_{s=1}^{S}\pi_s = 1$ from (4.1), the second term can be maximized by the method of Lagrange multipliers given $\boldsymbol{\beta}_s$. Set

$$L = \sum_{i=1}^{N}\sum_{s=1}^{S} h_i(s)\cdot \ln \pi_s - \lambda\left\{\sum_{s=1}^{S}\pi_s - 1\right\}.$$

Then we have $(S+1)$ set of equations by partially differentiating L with respect to π_s's and λ and setting the resulting formulas as zero as

$$\begin{cases} \frac{\partial L}{\partial \pi_1} = \frac{\sum_{i=1}^N h_i(1)}{\pi_1} - \lambda = 0, \\ \quad \vdots \\ \frac{\partial L}{\partial \pi_S} = \frac{\sum_{i=1}^N h_i(S)}{\pi_S} - \lambda = 0, \\ \frac{\partial L}{\partial \lambda} = -\sum_{s=1}^S \pi_s + 1 = 0. \end{cases} \tag{4.13}$$

From the first S equations in (4.13), we have

$$\lambda = \frac{\sum_{i=1}^N h_i(s)}{\pi_s}, \tag{4.14}$$

or

$$\pi_s = \frac{1}{\lambda} \sum_{i=1}^N h_i(s) \tag{4.15}$$

for $s = 1, \ldots, S$. Substitute these equations into the last equation in (4.13) to obtain

$$\frac{1}{\lambda} \sum_{i=1}^N h_i(1) + \cdots + \frac{1}{\lambda} \sum_{i=1}^N h_i(S) \quad = \quad 1$$

or

$$\sum_{i=1}^N (h_i(1) + \cdots + h_i(S)) \quad = \quad \lambda$$

or

$$N \quad = \quad \lambda,$$

since $h_i(1) + \cdots + h_i(S) = 1$. Therefore we have from (4.15)

$$\pi_s = \frac{\sum_{i=1}^N h_i(s)}{N} \tag{4.16}$$

for $s = 1, \ldots, S$. The solution (4.16) means that the relative size of segment s is the average of segment membership for s across all consumers in the sample.

Now from (4.12), $h_i(s) = \Pr(S_i = s|H_i; \boldsymbol{\beta}_s)$ can be calculated using the definition of conditional probability as[4]

$$h_i(s) = \frac{\Pr(S_i = s, H_i; \boldsymbol{\beta}_s)}{\Pr(H_i; \boldsymbol{\beta})} = \frac{\pi_s \cdot \Pr(H_i|S_i = s; \boldsymbol{\beta}_s)}{\sum_{s=1}^{S} \pi_s \cdot \Pr(H_i|S_i = s; \boldsymbol{\beta}_s)}. \tag{4.17}$$

By substituting (4.17) for (4.16), we obtain π_s.

As for the first term of the right hand side of (4.11) for segment s, the parameters can be estimated independently for each segment since the vectors of parameters $\boldsymbol{\beta}_s$ are independent across segments. Then the first term on the right hand side of (4.11) for segment s is written with the notation similar to (4.7) as

$$\begin{aligned} l_s(\boldsymbol{\beta}_s|\boldsymbol{H}) &= \sum_{i=1}^{N} h_i(s) \cdot \ln \Pr(H_i|S_i = s; \boldsymbol{\beta}_s) \\ &= \sum_{i=1}^{N} \sum_{t_i=1}^{T_i} \sum_{j=1}^{J} \{h_i(s) \cdot y_{ijt_i} \cdot \ln \Pr(Y_{it_i} = j|S_i = s; \boldsymbol{\beta}_s)\}. \end{aligned} \tag{4.18}$$

To implement EM algorithm, repeat the following steps.

EM algorithm

Step 0.1 Set $t = 0$. Set the initial values $\hat{\boldsymbol{\beta}}_s^{(0)}$ for $s = 1, \ldots, S$ and set $\pi_s^{(0)} = 1/S$ for $s = 1, \ldots, S$.

Step 0.2 Set $s = 1$. For $i = 1, \ldots, N$, calculate $h_i^{(t)}(s)$ by calculating $\Pr(Y_{it_i} = j|S_i = s; \boldsymbol{\beta}_s)$ using (4.3) first then (4.7) and (4.8) successively with $\hat{\boldsymbol{\beta}}_s^{(t)}$ and $\pi_s^{(t)}$ and substitute these interim results for (4.17). Set $s = s + 1$ and repeat **Step 0.2** until $s = S$.

[4]Note that $h_i(s)$ in (4.17) can be interpreted as the posterior distribution of consumer i's membership probability for segment s with prior distribution π_s and likelihood H_i given segment membership $S_i = s$ as we mentioned earlier.

Step 0.3 Calculate $E\left[l^{(t)}\left(\boldsymbol{\pi}^{(t)}, \hat{\boldsymbol{\beta}}^{(t)}|\mathbf{H}, \mathbf{Z}\right)\right]$ using (4.11).

Step 1 Set $s = 1$. Renew $\pi_s^{(t+1)}$ from (4.16) using $h_i^{(t)}(s)$.

Step 2 Estimate $\hat{\boldsymbol{\beta}}_s^{(t+1)}$ by maximizing (4.18) with (4.3) and $h_i^{(t)}(s)$ obtained previously. The actual maximization is done by the scoring or Newton-Raphson method.

Step 3 Renew $\Pr(Y_{it_i} = j|S_i = s; \boldsymbol{\beta}_s)^{(t+1)}$ by substituting $\hat{\boldsymbol{\beta}}_s^{(t+1)}$ obtained in **Step 2**.

Step 4 Calculate $h_i^{(t+1)}(s)$ from (4.17) with the renewed $\hat{\boldsymbol{\beta}}_s^{(t+1)}$ and $\pi_s^{(t+1)}$ for $i = 1, \ldots, N$. Set $s = s + 1$ and goto **Step 1**. If $s = S$, goto **Step 5**.

Step 5 Calculate $E\left[l^{(t+1)}\left(\boldsymbol{\pi}^{(t+1)}, \hat{\boldsymbol{\beta}}^{(t+1)}|\mathbf{H}, \mathbf{Z}\right)\right]$ using (4.11). If $E\left[l^{(t+1)}\left(\boldsymbol{\pi}^{(t+1)}, \hat{\boldsymbol{\beta}}_s^{(t+1)}|\mathbf{H}, \mathbf{Z}\right)\right]$ and $E\left[l^{(t)}\left(\boldsymbol{\pi}^{(t)}, \hat{\boldsymbol{\beta}}_s^{(t)}|\mathbf{H}, \mathbf{Z}\right)\right]$ are close enough, for example, less than small prescribed constant ϵ, stop the iteration as the expected log likelihood is maximized. Else set $s = 1$ and $t = t + 1$, and return to **Step 1**.

5 Empirical Results

We use ERIM database, the panel data of U.S. households in Sioux Falls, SD and Springfield, MO which was collected from 1st week of 1986 to 34th week of 1988. ERIM database is the data collected by the now-defunct ERIM division of A.C. Nielsen on panels of households in Sioux Falls and Springfield for academic research.[5]

We choose ketchup category for our empirical analysis for the following

[5]We acknowledge the James M. Kilts Center, University of Chicago Booth School of Business for letting us use the data.

reasons: First, since we are interested in consumers' brand choice behaviors with the possible presence of state dependence and/or variety-seeking behaviors, product categories in which a consumer exhibits strong genuine preference to specific brand are not suitable because a consumer would buy the specific brand anyway. Secondly, the products that are purchased with relatively high frequency are preferable, since we calibrate the effects of past brand choices on current purchasing occasion. In other words, the products that are purchased on irregular intervals would not suit as consumers may forget the brands they purchased on the previous occasion. After screening data, we have 137 households with 1,504 purchase records.

The summary statistics of the five SKUs analyzed in this study is listed in Table 5.1. "Coupon Usage" indicates the number of times coupon was used, and "Display" and "Feature" indicate the number of times they were promoted conditional on SKU being purchased. "Mean Value of Coupons" indicates the average value of coupons when they were used.

Table 5.1: Summary statistics of SKUs under study.

SKU	Share	Mean Price per oz.	Mean Value of Coupons	Coupon Usage	Display	Feature
Heinz 32 oz.	31.70%	3.37	1.24	37.94%	11.52%	43.82%
Heinz PLS 28 oz.	15.80%	4.38	2.41	33.09%	16.73%	34.55%
Hunt's PLS & GLS 32 oz.	14.30%	3.22	1.30	32.57%	11.93%	36.70%
Del Monte 32 oz.	6.40%	2.87	1.00	7.20%	11.20%	36.00%
Control 32 oz.	5.00%	2.65	1.64	3.77%	5.66%	24.53%

To calibrate the effectiveness our model, we tested the two other models; the model which only uses marketing variables as explanatory variable (Model 1); the model which incorporates GL variable along with the marketing variables (Model 2). The third model is our proposal model which incorporates GL variable and run in addition to marketing variables (Model 3). We have determined the number of segments based on AIC. The number of segments is chosen to be four because no significant increase in AIC is observed for Model 3 when the number of segments is increased from four to five as shown in Table 5.2. Comparing three different models with six segments, our proposal model has the lowest AIC value.

Table 5.2: AIC of the three models with different numbers of segments.

	Model 1	Model 2	Model 3
2 segments	1310.60	1058.19	1041.80
3 segments	948.52	861.27	835.64
4 segments	842.64	805.61	796.74
5 segments	819.09	796.54	794.21
6 segments	809.53	803.84	813.02

The estimated parameters of Model 3 are presented in Table 5.3. The coefficients of the brands indicate intrinsic preferences for the brands relative to Control 32 ounce which is used as the base brand. All the coefficients are consistent with our intuitions, i.e., all coefficients of price are negative, those of coupons are positive, and those of display and feature are positive in all segments. As for GL variable and run, they show interesting patterns which would have not been discovered if run was not included in the model. Only Segment 2 has an insignificant (absolute t-value less than 2) coefficient for run. The negative coefficients of run across three segments imply that the marginal utility of the same brand decreases as a result of repeated purchases of the same brand over time.

Now to see the behavioral patterns regulated by the combination of GL variable and run, we calculated the logit probabilities for each SKU and segment, assuming the situation where consumers repeatedly purchase the same brands five times in row. In the calculation, we used the average prices of SKUs and assuming no promotions took in place during the period. The results are shown in Table 5.4. For example, the number at $t = 3$ is the purchasing probability of the SKU being purchased given two consecutive purchases of that SKU. In Table 5.4, we have the information of state dependence or variety-seeking tendency of each segment for each SKU. Segment

Table 5.3: The parameters for the Model 3.

	segment 1	segment 2	segment 3	segment 4
Heinz 32 oz.	-0.076	4.279	1.885	1.437
	(0.0148)	(0.0183)	(0.0122)	(0.0155)
Heinz PLS 28 oz.	1.311	2.661	1.727	2.399
	(0.0072)	(0.0070)	(0.0065)	(0.0073)
Hunt's PLS & GLS 32 oz.	0.027	0.296	2.713	-2.530
	(0.0068)	(0.0050)	(0.0097)	(0.0055)
Del Monte 32 oz.	1.176	-2.141	1.071	-1.154
	(0.0071)	(0.0014)	(0.0043)	(0.0057)
Price	-0.848	-0.735	-0.666	-2.509
	(0.0701)	(0.0769)	(0.0682)	(0.0713)
Coupon	2.792	5.015	3.287	5.471
	(0.0211)	(0.0207)	(0.0183)	(0.0202)
Display	3.419	3.903	3.855	4.750
	(0.0070)	(0.0067)	(0.0067)	(0.0072)
Feature	5.622	2.504	2.938	5.894
	(0.0126)	(0.0121)	(0.0119)	(0.0127)
GL	4.303	0.743	1.737	5.607
	(0.0136)	(0.0115)	(0.0102)	(0.0126)
Run	0.505	*-0.127	-0.222	-0.123
	(0.0891)	(0.0663)	(0.0519)	(0.0529)
Size of Segments	0.261	0.234	0.289	0.217
Total Log Likelihood	-365.81			

* 90% level significance with t-value -1.917.

All the other coefficients were significant at the 0.05 level.

The numbers in parentheses are standard errors.

1 and 4 exhibit state dependence tendencies while segment 2 and 3 exhibit variety-seeking tendencies.

Table 5.4: Logit Probability of Purchase: Model 3

Heinz 32 oz.	Segment 1	Segment 2	Segment 3	Segment 4
t=0	8.8%	87.6%	20.6%	33.5%
t=1	56.1%	91.7%	42.4%	93.6%
t=2	74.9%	91.1%	40.4%	95.3%
t=3	86.7%	90.5%	37.7%	96.2%
t=4	93.1%	89.7%	34.7%	96.8%
t=5	96.4%	88.8%	31.3%	97.1%

Heinz PLS 28 oz.	Segment 1	Segment 2	Segment 3	Segment 4
t=0	15.0%	8.3%	9.0%	7.1%
t=1	70.1%	12.4%	21.9%	68.7%
t=2	84.5%	11.7%	20.5%	75.3%
t=3	92.3%	10.9%	18.8%	79.4%
t=4	96.1%	10.0%	16.8%	81.9%
t=5	98.0%	9.2%	14.8%	83.5%

Hunt's PLS & GLS 32 oz.	Segment 1	Segment 2	Segment 3	Segment 4
t=0	11.2%	1.8%	52.5%	0.9%
t=1	62.4%	2.8%	75.8%	21.6%
t=2	79.5%	2.7%	74.2%	27.6%
t=3	89.4%	2.5%	72.0%	32.6%
t=4	94.6%	2.3%	69.3%	36.3%
t=5	97.2%	2.0%	66.0%	38.8%

Del Monte 32 oz.	Segment 1	Segment 2	Segment 3	Segment 4
t=0	47.4%	0.2%	12.8%	9.0%
t=1	92.2%	0.3%	29.4%	74.0%
t=2	96.5%	0.3%	27.7%	79.8%
t=3	98.4%	0.3%	25.5%	83.3%
t=4	99.2%	0.3%	23.1%	85.5%
t=5	99.6%	0.2%	20.5%	86.7%

Control 32 oz.	Segment 1	Segment 2	Segment 3	Segment 4
t=0	17.6%	2.1%	5.1%	49.5%
t=1	73.9%	3.2%	13.2%	96.6%
t=2	86.9%	3.0%	12.3%	97.5%
t=3	93.5%	2.8%	11.1%	98.0%
t=4	96.8%	2.5%	9.9%	98.3%
t=5	98.3%	2.3%	8.6%	98.5%

The information in Table 5.4 can be used as a starting point for brand managers to plan their marketing strategies and promotional activities. For example, since consumers in segment 1 exhibit strong state dependence, Del

Monte may need to ensure it has enough amounts of promotions to retain consumers in this segment since Del Monte is preferred by segment 1 most. As consumers have low coefficients for coupon and display but have high coefficient for feature, Del Monte would want to increase feature to retain consumers from Segment 1.

6 Discussion

Overall, our model achieves the best AIC compared to the previously proposed models with a fair number of significant variables, indicating that the households are heterogeneous in their behavioral patterns over time. It gives important connotations for marketers, because the model without state dependence and variety seeking effects would be misleading in constructing the strategy and planning promotions as pointed out by previous research such as Keane (1997).

Unfortunately, the hybrid behavior was not detected in our analysis. This may be because most of consecutive purchases of the same SKU are three at most; about 90% of purchases in the data are shorter than three runs. Also, the products like ketchup, where the bottle is consumed through a relatively long period of time, a satiation effect may start during the consumption period and that may lead households to switch brand, i.e., the hybrid behavior is hidden as a result of the large package sizes of the ketchup. By using the products which are consumed in a relatively short period of time the hybrid behavior may has been detected.

For future researches, the model presented in this study can be tested using different data sets for the validity of the model. The new variable to explain state dependence and variety-seeking behaviors can be constructed as well. From a microeconomic perspective, the budget constraint can be incorporated in the model because households may switch brands depending on their budget at each shopping trip.

References

Bawa, Kapil (1990) 'Modeling inertia and variety seeking tendencies in brand choice behavior.' *Marketing Science* 9(3), 263–278

Chintagunta, Pradeep K. (1998) 'Inertia and variety seeking in a model of brand-purchase timing.' *Marketing Science* 17(3), 253–270

Givon, Moshe (1984) 'Variety seeking through brand switching.' *Marketing Science* 3(1), 1–22

Guadagni, Peter M., and John D. C. Little (1983) 'A logit model of brand choice calibrated on scanner data.' *Marketing Science* 2(3), 203–238

Kamakura, Wagner A., and Gary J. Russell (1989) 'A probabilistic choice model for market segmentation and elasticity structure.' *Journal of Marketing Research* 26(4), 379–390

Keane, Michael P. (1997) 'Modeling heterogeneity and state dependence in consumer choice behavior.' *Journal of Business and Economic Statistics* 15(3), 310–327

McAlister, Leigh, and Edgar Pessemier (1982) 'Variety seeking behavior: An interdisciplinary review.' *Journal of Consumer Research* 9(3), 311–322

Seetharaman, P.B. (2004) 'Modeling multiple sources of state dependence in random utility models: A distributed lag approach.' *Marketing Science* 23(2), 263–271

Train, K. (2003) *Discrete choice methods with simulation* (Cambridge Univ Pr)