An Analysis of Multicategory/brand Purchase Behavior with Hierarchical Bayes Multivariate Poisson Autoregressive Models

Miyazaki, Kei
Graduate School of Economics, Nagoya University
Furo-cho Chikusa-ku Nagoya-shi Aichi
464-8601 JAPAN
E-mail: miyazaki.behaviormetrics@gmail.com

1 Introduction

The main objective of marketing researchers is to understand and predict consumer purchase behavior, and because the time series data in which purchase histories are recorded for each consumer (that is, scanner panel data) are now easily obtainable, many studies have examined the longitudinal aspects of stability and changes in brand preferences or multicategory purchase behavior.

On a shopping trip, a household's incidence decisions are related across product categories because product categories serve as *substitutes* or *complements* in addressing the household's consumption needs. Brands in the same product category are regarded as substitutes, and well-known studies such as Guadagni and Little (1983) introduced the concept of the brand loyalty variable, and Kamakura and Russell (1989) and Bucklin and Gupta (1992) modeled consumer heterogeneity by using latent classes or finite mixtures. Regarding complementarity, Manchanda et al. (1999) employ the panel data multivariate probit (MVP) model in order to explain household-level contemporaneous incidence outcomes in multiple product categories.

Obviously, when purchase quantities of several brands or product categories are modeled, the time series variation of purchase quantities of substitute and complementary products should be considered. In previous studies, Bockenholt (1999) analyzed scanner panel data by using a latent class Poisson autoregressive model, but this model cannot deal with switching behavior across several brands because he considers only a single product category. Although Andrews and Currim (2009) modeled multiple categories/brands and purchase quantities, they do not consider the time series variation of purchase quantities. In particular, when brand choice behavior is considered, it is important to model the time series variation because switching behavior across brands is reflected in the time series variation of purchase quantities. Furthermore, the authors do not consider the difference of purchase behavior by consumer demographic characteristics. Song and Chintagunta (2007) and Mehta (2007) estimate a simultaneous model of the three purchase outcomes—incidence, brand choice, and quantity—in several product categories. However, these studies also do not consider either the time series variation of purchase quantities nor consumer demographic characteristics. Currently, few studies simultaneously model the time series variation of purchase quantities of substitute or complementary products and consumer demographic characteristics.

In this study, I develop a model that determines latent class membership via consumer demographic characteristics and identifies the causal relationships among purchase quantities of several product categories or brands for each latent class simultaneously. The proposed model can identify effective promotional activities for each consumer segment. For example, when family sizes are included in demographic variables, the following results may be obtained: The purchase quantity of product category A has significant predictive power for that of product category B in the segment to which nuclear households mainly belong, and on the other hand, the purchase quantity of category B has significant predictive power for that of category A in the segment to which bachelors mainly belong. That is, the product for which promotional activities should be conducted differs for each segment. Therefore, we consider the strategy of identifying the product that should be promoted by using the information obtained from the households' ID cards. Moreover, in the proposed model, promotional variables are included as the explanatory variables for the expectation of purchase quantities. Because the autoregression coefficients among product categories are estimated, this model setting enables identifying both the effective promotional activities for each consumer segment and the consumer's latent switching behavior across multiple products or brands when the promotional activities are not conducted.

2 Modeling Time Series Variation of Purchase Quantities

y_{itc} indicates the number of purchase quantity of the product category c for the i-th consumer in the time period t. We assume that y_{itc} follows Poisson distribution with the intensity parameter μ_{itc}. Thus, the probability of $y_{itc} = p$ is expressed as follows:

$$(1) \quad p(y_{itc} = p|\mu_{itc}) = \frac{e^{-\mu_{itc}}\mu_{itc}^{y_{itc}}}{y_{itc}!}.$$

When we consider the purchase behaviors of several products, there exist interrelated influences among purchase quantities in the case of both complementary and substitute products (Manchanda et al., 1999). This can be modeled by assuming correlation relationships between the intensity parameters of Poisson distributions.

The Expression of Intensity Parameters

The proposed model assumes autoregressive relationships for the intensity parameter vector $\boldsymbol{\mu}_{it} = (\mu_{it1}, \cdots, \mu_{itC})'$ that determines purchase quantity as in Eq. (1). The vector autoregressive model is assumed for the logarithmically transformed vector $\boldsymbol{\eta}_{it} = (\eta_{it1}, \cdots, \eta_{itC}) = (\log \mu_{it1}, \cdots, \log \mu_{itC})'$. The reason for transforming $\boldsymbol{\mu}_{it}$ is that $\boldsymbol{\mu}_{it}$ is always positive. Let $j (j = 1, \cdots, M)$ be the lag in autoregressive models and $\boldsymbol{\eta}_{i[t-j]} = (\eta_{i[t-j]1}, \cdots, \eta_{i[t-j]C})'$ be the i-th consumer's log intensity parameter vector in the time period $t - j$. X_{it} denotes the marketing variables for the i-th consumer. X_{it} includes the data about price, display, and feature. We consider that the consumers are separated into heterogeneous subpopulations; that is, we assume the latent classes $(l = 1, \cdots, L)$. k_i denote the latent class that the i-th consumer belongs to. Let $\boldsymbol{\beta}_l$ be the regression coefficient vector of the marketing variable X_{it} in the l-th latent class. We assume that the autoregressive coefficient matrix also differs across classes. $\gamma_{lc_1c_2j}$ is the autoregression coefficient with the l-th latent class and the j-th lag when the dependent and explanatory variables are categories c_1 and c_2 respectively.

Under the above notation, the intensity parameter vector $\boldsymbol{\eta}_{it} = (\eta_{it1}, \cdots, \eta_{itC})'$ is defined as follows:

$$(2) \quad \boldsymbol{\eta}_{it} = X_{it}\boldsymbol{\beta}_{k_i} + \sum_{j=1}^{M} \Gamma_{k_i \cdot\cdot j}\boldsymbol{\eta}_{i[t-j]} + \boldsymbol{\epsilon}_{it}, \quad \boldsymbol{\epsilon}_{it} \sim N(\mathbf{0}, \Sigma_l),$$

where

$$(3) \quad \Gamma_{k_i \cdot\cdot j} = \begin{pmatrix} \gamma_{k_i 11j} & \cdots & \gamma_{k_i 1Cj} \\ \vdots & \ddots & \vdots \\ \gamma_{k_i C1j} & \cdots & \gamma_{k_i CCj} \end{pmatrix}.$$

X_{it} includes intercepts, as shown below:

$$(4) \quad X_{it} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & p_{it1} & d_{it1} & f_{it1} \\ 0 & 1 & 0 & \cdots & 0 & p_{it2} & d_{it2} & f_{it2} \\ & & \vdots & & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & p_{itC} & d_{itC} & f_{itC} \end{pmatrix}.$$

In this model, we do not assume that the previous marketing variable $X_{i(t-j)}$ influences the current intensity parameter $\boldsymbol{\eta}_{it}$.

Determining Latent Class Membership

In marketing science, it is usual to consider that latent class membership is determined by consumer demographic characteristics (Gupta and Chintagunta, 1994). Therefore, we assume that k_i, the indicator variable of latent class membership, follows nominal logistic distribution in which demographic variables are the explanatory variables:

$$(5) \quad p(k_i = l|\boldsymbol{\pi}, \boldsymbol{z}_i) = \frac{\exp(\boldsymbol{\pi}'_l \boldsymbol{z}_i)}{\sum_{j=1}^{L} \exp(\boldsymbol{\pi}'_j \boldsymbol{z}_i)},$$

where $\boldsymbol{\pi} = (\boldsymbol{\pi}'_1, \cdots, \boldsymbol{\pi}'_L)'$ is the regression coefficient of the demographic variable, but for identifiability of the model, we assume $\boldsymbol{\pi}_L = \boldsymbol{0}$.

Since the proposed model includes discrete variables (demographic variables and purchase quantity data) and latent variables (intensity parameter and latent class indicator variable), it is difficult to apply the maximum likelihood estimation method to this model because many integrals must be calculated. Thus, in this study, we describe the model as a hierarchical Bayes model and estimate the parameters by using the Markov chain Monte Carlo algorithm.

3 Real Data Analysis: IRI Marketing Data Set

We illustrated the applicability of our method, using a small part of the IRI marketing data set (Bronnenberg, Kruger and Mela, 2008). The data span a period of 12 months from January 2002 to December 2002. We analyzed the data on 3 main frozen food brands (Conagra, Heinz, and Nestle as an example of a data set for substitute products, and the data on toothbrushes and toothpaste, for complementary products. The span of 1 lag is 1 month. The demographic variables we used are combined pre-tax income of the household, family size, and purchase probability per month. For each product category, the data set includes the marketing variables price, display, and feature. Households that had made at least one purchase in any of the product categories during the 12-month period were selected. This resulted in a sample of 6,070 households. We attempted to construct X_{it} by identifying the marketing variables that were marked in the store visited by person i at the time of his/her visit in month t, but because only the marketing variables of the products actually purchased are recorded in the IRI scanner panel data, the marketing variables of non-purchased products are missing. To fill in the missing data we used the average marketing variables marked in that store during month t.

An Analysis of Substitute Products Data—Frozen Food Data

As an example of the analysis of data for substitute products, we prepared the panel data of the top 3 brands in the frozen food market in 2002. This analysis shows that the proposed model can be used for understanding brand switching behavior.

The descriptive statistics for the above 3 brands are listed in Table 1. Mean purchase quantity is the value that is obtained by averaging the purchase quantities per month for all the households.

In order to determine the most appropriate number of latent classes and lags, we set 20 models in which the numbers of classes and lags are from 1 to 5 and from 1 to 4 respectively. We transformed the price values per 16 oz logarithmically. After employing 5,000 burn-in iterations, we employed 5,000 MCMC iterations in order to calculate BIC. The results indicate that the model with 4 classes and 1 lag is the most appropriate.

The sizes of the 4 classes are 54.6%, 13.8%, 16.6%, and 15.0%, respectively. The regression coefficients of demographic variables for latent class membership are given in Table 2. The regression coefficients

Table 1: Descriptive Statistics of Marketing Variables (Substitute Products)

Brand Name	Mean Purchase Quantity	Price ($/16 oz)	Display	Feature
Conagra	6.206	3.461	0.1504	0.05187
Heinz	4.528	4.189	0.1903	0.04712
Nestle	7.402	4.233	0.1520	0.03844

of marketing variables for expectation of purchase quantities (that is, intensity parameters) are listed in Table 3, and autoregression coefficients, in Table 4. Dependent variables in Eq. (2) are listed in the rows of the table, and explanatory variables, in the columns. For example, the value of autoregression coefficient of "Heinz $\leftarrow$ Conagra" in class 1 is -0.3992.

The households that belong to class 1 have the highest pre-tax income, largest family size, and lowest purchase probability; that is, the households in this class are apt to buy all their needs at one time. In Table 3, all the intercept values are significantly negative, and all the regression coefficients of marketing variables are significant; that is, this segment has potential for the effect of promotional activities. Table 4 shows that the regression coefficients of own brands are significantly positive, except for Nestle This result implies that this class is apt to switch to other brands from Nestle Another feature is that the more the products of Conagra are purchased during month $t-1$, the less the products of Heinz are purchased during month t. This result means that in this class, once consumers switch to Conagra, they find it difficult to return to Heinz. Table 3 shows that in class 2, only the intercept value of Nestleis significant, and the conduction of both feature and display does not heighten buying motivation. However, in Table 4, autoregression coefficients for own brands are high, which indicates that the consumers in this class buy the same brand over time "out of habit," and the regression coefficient from Nestleto Heinz is significantly negative, which means that in class 2, Nestlecan successfully grab a share of the market from Heinz. Class 3 is similar to class 2 in demographic characteristics, but the features of the regression coefficients of marketing variables and intercepts are clearer here than class 2. In particular, sensitivity to price values is higher. In Table 4, the presence of many significant values indicates that the consumers in this class are apt to have variety-seeking behavior. As well as classes 2 and 3, pre-tax income of class 4 is lower than that of class 1 and family size is also smaller, but the regression coefficients of marketing variables are all significant; that is, this class has potential for the effect of promotional activities. Table 4 shows that switching behavior exists between Conagra and Heinz, and Conagra and Nestle These pairs are scrambling for their market shares.

Table 2: Regression Coefficients for Demographic Variables (Substitute Data)

	Class 1	Class 2	Class 3	Class 4
Combined Pre-Tax Income	**0.3460** (0.2394,0.4589)	-0.07098 (-0.1856,0.04135)	-0.07853 (-0.1987,0.03718)	0
Family Size	**0.6219** (0.3416,0.9299)	0.07887 (-0.2407,0.4054)	0.02138 (-0.3041,0.3673)	0
Purchase Probability per Month	**-4.944** (-6.181,-3.721)	0.5114 (-0.4974,1.546)	0.8609 (-0.1662,1.911)	0

(parentheses: 95 % Bayesian credible intervals)

Table 3: Intercepts and Regression Coefficients of Marketing Variables (Substitute Products)

	Class 1	Class 2	Class 3	Class 4
Conagra	**-0.9075**	0.04645	0.1235	**-0.7289**
	(-1.122,-0.6934)	(-0.01605,0.1097)	(-0.04602,0.2861)	(-0.9992,-0.4645)
Heinz	**-1.142**	0.02538	0.03590	**-1.135**
	(-1.374,-0.9236)	(-0.04859,0.1081)	(-0.1889,0.2355)	(-1.443,-0.8077)
Nestle	**-0.9496**	**0.1010**	**0.2988**	-0.2443
	(-1.214,-0.6807)	(0.01197,0.1897)	(0.1366,0.4629)	(-0.5257,0.02492)
Price	**-0.3524**	**-0.02861**	**-0.3413**	**-0.3438**
	(-0.4157,-0.2843)	(-0.05322,-0.002964)	(-0.4015,-0.2768)	(-0.4112,-0.2676)
Feature	**0.6190**	0.03171	-0.1580	**1.071**
	(0.3066,0.9288)	(-0.06622,0.1384)	(-0.4236,0.1070)	(0.7266,1.428)
Display	**1.327**	-0.08532	0.3582	**1.293**
	(0.9911,1.666)	(-0.2581,0.05497)	(-0.08283,0.8064)	(0.7955,1.743)

(parentheses: 95 % Bayesian credible intervals)

5 Discussion

In this study, I proposed a model for examining the time series variation of purchase quantities of several product categories or brands for each segment that differs in consumer demographic characteristics. The proposed model includes marketing variables as the explanatory variables. Therefore, this model enables identifying autoregressive relationships among purchase quantities of several product categories when the effects of marketing variables are fixed. When brands are the dependent variables, we can analyze brand switching or brand loyalty, and when the product categories are the dependent variables, we can analyze the latent purchase cycle among the product categories.

In this study, we did not consider time series variations and the influence of promotional activities on price elasticity, but according to Mela, Gupta, and Lehmann (1997), advertisements have long-term effects and while advertisements decrease price elasticity, in the long-term, price promotions increase price elasticity. Therefore, it is possible that switching relations among several brands change owing to promotional activities such as advertisements over the long term. In order to study this, it is necessary to develop a model in which the regression coefficients of promotions change with time. This topic could be addressed in future research.

REFERENCES

Andrews, R.L., & Currim, I.S. (2009). Multi-stage purchase decision models: Accommodating response heterogeneity, common demand shocks, and endogeneity using disaggregate data. *International Journal of Research in Marketing*, 26, 197-206.

Bockenholt, U. (1999). Mixed INAR(1) Poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data. *Journal of Econometrics*, 89, 317-338.

Bronnenberg, B.J., Kruger, M.W. and Mela, C.F. (2008), "Database paper: The IRI marketing data set," *Marketing Science*, 27(4), 745-748.

Bucklin, R.E. and Gupta, S. (1992), "Brand Choice, Purchase Incidence, and Segmentation: An Integrated Approach," *Journal of Marketing Research*, 29, 201-215.

Table 4: Autoregression Coefficients between Intensity Parameters (Substitute Data)

		Conagra	Heinz	Nestle
Class 1	Conagra	**0.3152**	-0.06205	-0.2104
		(0.1284,0.4792)	(-0.3113,0.1364)	(-0.5176,0.1554)
	Heinz	**-0.3992**	**0.4926**	-0.2133
		(-0.6541,-0.1218)	(0.2970,0.6836)	(-0.6981,0.1801)
	Nestle	-0.09030	0.09815	-0.1267
		(-0.2761,0.1254)	(-0.1258,0.3637)	(-0.4467,0.1762)
Class 2	Conagra	**1.157**	-0.008232	-0.03428
		(1.045,1.309)	(-0.1092,0.1358)	(-0.1235,0.04280)
	Heinz	-0.05449	**1.096**	**-0.1222**
		(-0.1899,0.1021)	(0.9454,1.289)	(-0.2595,-0.008723)
	Nestle	-0.001044	0.006437	**0.9523**
		(-0.09325,0.08474)	(-0.07304,0.08846)	(0.8898,1.014)
Class 3	Conagra	**0.5901**	**-0.5363**	-0.1091
		(0.4510,0.6979)	(-0.6362,-0.4439)	(-0.2567,0.01192)
	Heinz	**-0.4161**	**0.2459**	**-0.1759**
		(-0.5451,-0.2967)	(0.1272,0.3632)	(-0.3534,-0.02144)
	Nestle	**-0.3458**	**-0.5361**	**0.7009**
		(-0.4512,-0.2508)	(-0.6409,-0.4391)	(0.5574,0.8308)
Class 4	Conagra	**0.2229**	**-0.2574**	**-0.5537**
		(0.07458,0.3538)	(-0.4210,-0.07194)	(-0.6794,-0.4287)
	Heinz	**-0.3953**	-0.1179	-0.2163
		(-0.6283,-0.1536)	(-0.3767,0.1547)	(-0.4170,0.007160)
	Nestle	**-0.6680**	0.05072	**0.1710**
		(-0.8103,-0.5363)	(-0.09588,0.2136)	(0.02175,0.2970)

(parentheses: 95 % Bayesian credible intervals)

Guadagni, P. and Little, J.D.C. (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2(3), 203-238.

Gupta, S. and Chintagunta, P.K. (1994), "On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models," *Journal of Marketing Research*, 31(1), 128-136.

Kamakura, W.A. and Russell, G.J. (1989), "A Probabilistic Choice Model of Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26, 379-390.

Manchanda, P., Ansari, A. and Gupta, S. (1999), "The "Shopping Basket": A Model for Multicategory Purchase Incidence Decisions," *Marketing Science*, 18(2), 95-114.

Mehta, N. (2007). Investigating consumers ' purchase incidence and brand choice decisions across multiple product categories: A theoretical and empirical analysis. *Marketing Science*, 26(2), 196-217.

Mela C.F., Gupta S. and Lehmann, D.R. (1997), "The Long-term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal of Marketing Research*, 34(2), 248-261.

Song, I., & Chintagunta, P.K. (2007). A discrete-continuous model for multicategory purchase behavior of households. *Journal of Marketing Research*, 44(4), 595-612.