AN INTEGRATED PURCHASE MODEL USING GAUSSIAN COPULA

Genya Kobayashi*, Kazuhiko Kakamu**, Eisaku Sato**, and Hideo Kozumi***

This paper proposes an integrated purchase model of household category purchase incidence, brand choice, and purchase quantity choice using the Gaussian copula. In contrast to the existing model, we assume the general form of the dependence parameter matrix for Gaussian copula. The proposed approach allows us to decompose the joint probability of the purchase outcomes into the conditional probability of one decision given the others. The price elasticities derived based on these conditional probabilities can fully reflect the underlying dependence among the decisions. The conditional probabilities are also utilized to predict future responses mimicking the sequence of the purchase decisions. The proposed model is applied to scanner panel data for the dishwashing soap category. We find that there exists a very strong positive dependence between incidence and brand choice, while dependence between incidence and quantity choice, and between brand choice and quantity choice is negative. The main sources of the overall behavioral response to price are found to be the incidence and brand choice decisions, while the quantity choice decision is hardly influenced by price change after decisions are made on the category of purchase and brand choice.

1. Introduction

In the marketing literature, the household purchase decisions on a shopping occasion are typically considered to consist of the following: whether or not to buy from the product category (incidence), which brand of the product to buy (brand choice), and how many units of the product to buy (quantity). These purchase decisions can be influenced by marketing mix variables. Understanding how households respond to them in each decision is crucial for retailers and manufacturers to effectively implement marketing actions.

When the household purchase behavior is considered, these three decisions must be taken into account simultaneously. Since the decisions are not made independently, there exists unobserved interdependence among the decisions. Therefore, it would be infeasible to treat each decision separately or to take a partial approach that ignores one of the decisions. Studies that do the former include Gupta (1988), while studies that do the latter include Bucklin and Lattin (1991), Mela et al. (1998), and Jedidi et al. (1999); which considered incidence and brand choice, incidence and quantity, and brand choice and quantity, respectively. Estimates based on such models can suffer from selectivity bias.

A number of integrated purchase models that analyze the three decisions simul-
taneously have been proposed in the marketing literature. Early integrated models include Chiang (1991) and Chintagunta (1993), where the optimal outcomes of incidence, brand choice and quantity are derived from the utility maximization. They modeled quantity choice by simple regression models. However, since households cannot buy products in quantities that are not multiple of available package sizes, quantity outcomes are observed as discrete counts. Therefore, their regression approach which can produce negative and continuous outcomes is not appropriate. The model of Bucklin et al. (1998) modeled incidence and brand choice through the nested logit model and quantity choice was modeled through the truncated Poisson regression model. Their truncated Poisson regression model is more reasonable for quantity choice since positive and discrete outcomes are guaranteed. They further introduced household segmentation using latent mixtures. Using the data for the paper towel category, the sales responses of households on the three decisions in one segment were found to be quite different from those in another segment. Their model was reconsidered in Ailawadi et al. (2007) who investigated the impact of household stockpiling on sales of perishable goods. More recently, more integrated purchase models have been proposed in the contexts of the multi-category demand (Song and Chintagunta, 2006), multi-store and aggregate demand (Nair, et al., 2005), and dynamic purchase behavior (Erdem et al., 2003; Sun, 2005; Chan et al., 2006).

In the household purchase behavior analysis, interdependence among the decisions also plays an important role, but it has often been ignored. Dependence measures can provide insight into how a change in one decision influences another. For example, if there is a positive dependence between the incidence and choice of one brand, purchases from the category are associated with this brand, which may indicate household favoritism of for that brand. Similarly, if there is a positive dependence between incidence and quantity, the category purchases are associated with large purchase quantities. One attempt to quantify the dependences can be found in Chib et al. (2004), which introduced the non-purchase option into the multinomial probit model for the brand choice decision and allowed the unobserved factors of the choices to be correlated. This model was shown to have a general dependence structure since all the dependences between the purchase incidence and choice of each brand can be quantified. However, the majority of studies introduced correlation structures merely to avoid selectivity biases and measuring interdependence has rarely been considered.

Zhang et al. (2005) first focused on how one decision depends on the others in an integrated model based on a model specification similar to that of Bell et al. (1999). Their unique modeling approach is based on the Gaussian copula (see, e.g., Nelsen, 2006; Trivedi and Zimmer, 2007), although it was not explicitly stated. Zhang et al. (2005) modeled the decisions following Bucklin et al. (1998) and expressed the joint probability of the three outcomes using the Gaussian copula. The interdependence among the decisions can be measured through the copula parameters. The model was further used by Zhang et al. (2012) to investigate the influence of promotions on purchase decisions when consumers infer future deals. Ebling and Klapper (2010) compared the performance of the models of Zhang et al. (2005) and
Song and Chintagunta (2006) which they called the unified model. The approach of Zhang et al. (2005) has been reported to be particularly useful when household behavior differs across the three decisions.

The flexibility of the copula modeling approach has not been fully exploited yet as far as we know. Zhang et al. (2005) ruled out unobserved incidence–choice dependence; that is, they set the parameters of the copula for incidence–choice dependence to zero. Instead, they used the nested logit model to combine the incidence and brand choice models and tried to capture incidence–choice dependence through the inclusive value measure. An implication is that incidence and brand choice are treated independently in the copula modeling framework. Since purchase decisions are made in some order, inference using an integrated model should be able to mimic the decision process by recovering the conditional probabilities of each decision given the others.

This paper proposes an alternative model specification for the integrated purchase model using the Gaussian copula and presents a set of inference procedures that takes advantage of the flexibility of the proposed modeling approach. We model the joint probability of incidence, brand choice, and quantity outcomes using the Gaussian copula with the general dependence structure. The proposed approach, which represents the joint probability of the outcomes, is feasible since the outcomes are observed only jointly even though the decisions are made in some order. We then show that the joint probability can be decomposed into the conditional probabilities of each decision given the others. Particularly following the literature on integrated purchase models, which include Gupta (1988), Bucklin et al. (1998), and Ailawadi et al. (2007), we consider the brand choice probability conditionally on incidence and quantity choice probability conditionally on the incidence and brand choice. We will also show that these conditional probabilities can be utilized in the following two ways. First, they are used to derive the price elasticities of the decisions conditionally on the previous decisions. The decomposition of overall elasticity into such conditional elasticities is often of interest in this context (see, e.g., Gupta, 1988). On the other hand, since in Zhang et al. (2005) incidence and brand choice were treated as independent, decomposition as in our approach was not available, and thus, the elasticities were derived from the marginal models. The interdependence, thus, was not reflected in their price elasticities. Second, the conditional probabilities can be also used to predict future outcomes in accordance with the sequence of the purchase decisions.

The rest of this paper is organized as follows. In Section 2, we introduce the integrated purchase model based on the Gaussian copula with the general dependence structure. Then, we derive the conditional price elasticities and introduce the method for predicting future outcomes using the conditional probabilities of the decisions obtained from the Gaussian copula. In Section 3, the proposed model is applied to scanner panel data for the dishwashing soap category. Finally, we conclude our discussion in Section 4.
2. Proposed Approach

2.1 Copula Modeling Approach

The purchase outcomes made by households are observed only jointly. In other words, we only know whether the category purchase occurred first and which brand and how many units were chosen after that. However, we do not have information on the order in which the decisions were made. When we predict future outcomes or compute elasticities, it is important to recover the decision order since it is known that the decision order depends on the context or task of the shopping occasion, and it may influence purchase outcomes (e.g., Nowlis et al., 2010). Therefore, to derive further insights from the data, a statistical model should be able to model the purchase outcomes jointly and to decompose them into a sequence of models for each decision conditionally on the others. Furthermore, since there exists some interdependence among the decisions, a model that explicitly incorporates the unobserved interdependence would be desirable. The proposed copula modeling approach conveniently satisfies these requirements.

Jointly considering the outcomes of incidence, brand choice, and quantity means that we are dealing with a trivariate joint distribution. Despite recent demands for multivariate analysis, it is generally difficult to find a multivariate counterpart of an univariate distribution. This is particularly the case when the marginal distributions are non-normal or discrete (for example, the Poisson distribution). Recently, copulas have drawn substantial attention as a method for combining marginal distributions to construct a multivariate distribution. Copulas have been applied in the fields of finance (Cherubini et al., 2004) and economics (Zimmer and Trivedi, 2006), among others. The usefulness of the copula approach has been also recognized in the marketing literature (Danaher and Smith, 2011).

The popularity of the copula approach comes from Sklar’s theorem (see e.g., Sklar, 1959; Nelsen, 2006). Suppose the random variables $X_1, \ldots, X_n$ have the distribution functions $F_1, \ldots, F_n$, respectively, and have the joint distribution $F$. Sklar’s theorem states that there exists an $n$-dimensional copula $C : [0, 1]^n \rightarrow [0, 1]$ such that

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \quad (2.1.1)$$

Conversely, if $C$ is an $n$ dimensional copula and $F_1, \ldots, F_n$ are the distribution functions, the function $F$ in (2.1.1) is an $n$-dimensional distribution function with margins $F_1, \ldots, F_n$. That is, once the marginal models are specified, the joint distribution can be constructed using the chosen copula regardless of the complexity of the margins. Therefore, it is possible to model each decision in detail without complicating the modeling of the joint distribution.

2.2 Gaussian Copula

This paper focuses on the Gaussian copula (see Song, 2000). The three dimensional
Gaussian copula \((n = 3)\), which is of our particular interest, is given by
\[
C(u_1, u_2, u_3) = \Phi_3 \left( \Phi^{-1}(u_1), \Phi^{-1}(u_2), \Phi^{-1}(u_3); R \right),
\]
where \(u_i = F_i(x_i), \ i = 1, 2, 3\), \(\Phi^{-1}(\cdot)\) is the inverse of the distribution function of the univariate standard normal distribution \(\Phi(\cdot)\), \(\Phi_3(\cdot, \cdot, \cdot; R)\) is the distribution function of the trivariate normal distribution \(\mathcal{N}(0, R)\), and \(R\) is the matrix of the dependence parameters given by
\[
R = \begin{pmatrix}
1 & \rho_{12} & \rho_{13} \\
\rho_{12} & 1 & \rho_{23} \\
\rho_{13} & \rho_{23} & 1 \\
\end{pmatrix}, \quad -1 < \rho_{12}, \rho_{13}, \rho_{23} < 1.
\]

Note that the interpretation of the dependence parameters differs from that of the correlation matrix of the normal distribution since the margins are typically not normal. Instead, to quantify the dependence, Spearman’s rho and Kendal’s tau, which can be easily computed from the Gaussian copula parameters, can be used.

The Gaussian copula is useful for modeling household purchase behavior. While some copulas can only handle either positive or negative dependence, the Gaussian copula is able to capture both positive and negative dependences. If \(\rho_{kl}\) is positive (negative), \(X_k\) and \(X_l\) have positive (negative) dependence, and if \(\rho_{kl} = 0\), \(X_k\) and \(X_l\) are independent. Moreover, while most copulas can handle only bivariate distributions and multivariate extension is generally extremely difficult even for trivariate cases (see, e.g., Zimmer and Triverdi, 2006), an extension to the \(n\) dimensional Gaussian copula is straightforward.

An advantage of the proposed approach is that we can obtain the decomposition of the joint probability into the sequence of the conditional probabilities by exploiting the structure of the Gaussian copula. For example, the conditional distribution function of \(X_2\) given \(X_1\) can be written as
\[
F_{2|1}(x_2|x_1) = \frac{\Phi_2(x_1^*, x_2^*)}{\Phi(x_1^*)} = \Phi \left( \frac{x_2^* - \rho_{12} x_1^*}{\sqrt{1 - \rho_{12}^2}} \right),
\]
where \(x_i^* = \Phi^{-1}(u_i), \ i = 1, 2\). Thus, the conditional distribution function of \(X_2\) given \(X_1\) can be written as the distribution function of \(\mathcal{N}(\rho_{12} x_1^*, 1 - \rho_{12}^2)\). Expression (2.2.4) corresponds to the conditional probability of choosing a brand given the purchase incidence. Similarly, the distribution function of \(X_3\) conditional on \(X_1\) and \(X_2\) is given by
\[
F_{3|12}(x_3|x_1, x_2) = \Phi \left( \frac{x_3^* - r_{21} R_{11}^{-1} x_{12}^*}{\sqrt{1 - r_{21} R_{11}^{-1} r_{12}}} \right),
\]
where \(x_{12}^* = (x_1^*, x_2^*)', R_{11}, r_{12}\), and \(r_{21}\) are the partitions of \(R\) in (2.2.3) given by
\[
R_{11} = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}, \quad r_{12} = r_{21}' = (\rho_{13}, \rho_{23}).
\]
Expression (2.2.5), in turn, corresponds to the conditional probability of the quantity given the incidence and brand choice. We will show how these conditional probabilities are utilized in making inferences for the integrated purchase model in the next subsection.

2.3 Integrated Purchase Model

Now we introduce our integrated purchase model which expresses the joint probability of the purchase outcomes using the Gaussian copula. For the households \( h = 1, \ldots, H \) observed over shopping occasions \( t = 1, \ldots, T_h \), let \( y_{ht} \) denote the indicator of incidence which equals 1 if a purchase from the category occurs and 0 otherwise, let \( b_{ht} \in \{1, \ldots, J\} \) denote the index of the chosen brand, and let \( q_{ht} \) denote the purchase quantity. Assuming that the incidence, brand choice, and quantity outcomes are simultaneously realized from some joint distribution, we consider the joint probability given by

\[
P_{htj} = \text{Prob}(y_{ht} = 1, b_{ht} = j, q = q_{ht}) \quad (2.3.7)
\]

for \( j = 1, \ldots, J \). Following Zhang et al. (2005), the joint probability is given by

\[
P_{htj} = \Phi_3 \left( \Phi^{-1}(Y_{ht}), \Phi^{-1}(B_{htj}), \Phi^{-1}(\text{Prob}(q \leq q_{ht})); R^j \right)
- \Phi_3 \left( \Phi^{-1}(Y_{ht}), \Phi^{-1}(B_{htj}), \Phi^{-1}(\text{Prob}(q \leq q_{ht} - 1)); R^j \right), \quad (2.3.8)
\]

where \( Y_{ht} = \text{Prob}(y_{ht} = 1) \), \( B_{htj} = \text{Prob}(b_{ht} = j) \) and \( R^j \) is the matrix of the dependence parameters given by

\[
R^j = \begin{pmatrix}
1 & \rho_{12}^j & \rho_{13}^j \\
\rho_{12}^j & 1 & \rho_{23}^j \\
\rho_{13}^j & \rho_{23}^j & 1
\end{pmatrix}. \quad (2.3.9)
\]

The matrix of the dependence parameters \( R^j \) plays an important role in the proposed model framework. The dependences between the incidence and choice of brand \( j \), between the incidence and quantity, and between the quantity and choice of brand \( j \) are captured by \( \rho_{12}^j \), \( \rho_{13}^j \), and \( \rho_{23}^j \), respectively. For example, a positive \( \rho_{12}^j \) implies that the purchases from the category tend to result in the choice of brand \( j \). Similarly, if \( \rho_{13}^j \) is positive, it implies that the category purchases are associated with large purchase quantities. This may, in turn, imply that purchases from the category induce households stockpiling of the products.

The existing marketing literature predominantly assumes that the brand choice decision is made conditionally on the incidence decision and that the quantity is determined conditionally on the incidence and brand choice decisions (e.g., Bucklin et al., 1998; Ailawadi et al., 2007). We also adopt this decision order in our model. Then, the inference is based on the following conditional probabilities which are obtained by decomposing the joint probability (2.3.7) as
\[ \text{Prob}(y_{ht} = 1, b_{ht} = j, q = q_{ht}) = \text{Prob}(y_{ht} = 1) \text{Prob}(b_{ht} = j \mid y_{ht} = 1) \text{Prob}(q = q_{ht} \mid y_{ht} = 1, b_{ht} = j). \] (2.3.10)

It is necessary for us to be able to recover the conditional probabilities of the decisions from (2.3.8). Using the results in (2.2.4) and (2.2.5), they are obtained as

\[ \text{Prob}(b_{ht} = j \mid y_{ht} = 1) = \Phi \left( \frac{\Phi^{-1}(B_{htj}) - \rho_{j12} \Phi^{-1}(Y_{ht})}{\sqrt{1 - (\rho_{j12})^2}} \right) \] (2.3.11)

and

\[ \text{Prob}(q = q_{ht} \mid y_{ht} = 1, b_{ht} = j) = \Phi \left( \frac{\Phi^{-1}(\text{Prob}(q \leq q_{ht})) - r_{j21}(R_{11}^j)^{-1}x^*}{\sqrt{1 - r_{j21}(R_{11}^j)^{-1}r_{j12}}} \right) - \Phi \left( \frac{\Phi^{-1}(\text{Prob}(q \leq q_{ht} - 1)) - r_{j21}(R_{11}^j)^{-1}x^*}{\sqrt{1 - r_{j21}(R_{11}^j)^{-1}r_{j12}}} \right), \] (2.3.12)

respectively, where \( x^* = (\Phi^{-1}(Y_{ht}), \Phi^{-1}(B_{htj}))' \) and the partitions \( R_{11}^j, r_{12}^j, r_{21}^j \) of (2.3.9) are as in (2.2.6). We utilize these conditional probabilities in Section 3.4 for price elasticities and in Section 2.7 for prediction.

Our model can be considered a generalization of that of Zhang et al. (2005). They restricted the parameters of the Gaussian copula as

\[ \tilde{R}^j = \begin{pmatrix} 1 & 0 & \rho_{13} \\ 0 & 1 & \rho_{23}^j \\ \rho_{13} & \rho_{23}^j & 1 \end{pmatrix}. \] (2.3.13)

Since incidence and choice were treated as if they were independent in the copula, it is impossible to capture incidence–choice dependence under the specification of \( \tilde{R}^j \) (Chib et al., 2004).

### 2.4 Marginal Models

#### 2.4.1 Purchase Incidence

We describe the model for the purchase incidence decision. The model for incidence is essentially the random coefficient binary probit model. The incidence probability is given by

\[ Y_{ht} = \text{Prob}(z_{ht} > 0), \] (2.4.14)

where \( z_{ht} \) is the unobserved utility of household \( h \) at shopping occasion \( t \) for the
category purchase. We assume that \( z_{ht} \) is given by

\[
z_{ht} = \gamma_{h1} + \gamma_{h2} CR_h + \gamma_{h3} INV_h + \sum_{j=1}^{J} \gamma_{h,3+j} \text{PRICE}_{htj} + \sum_{j=1}^{J} \gamma_{h,3+J+j} \text{FEATURE}_{htj} + \epsilon_{ht},
\]

(2.4.15)

where \( INV_h \) is the inventory level of household \( h \) at shopping occasion \( t \), \( CR_h \) is the time invariant household specific consumption rate, \( \text{PRICE}_{htj} \) is the unit price of brand \( j \), and \( \text{FEATURE}_{htj} \) takes value 1 if the brand \( j \) is promoted and 0 otherwise.

We computed \( INV_h \) from the equation

\[
INV_h = INV_{h,t-1} + q_{h,t-1} - CR_h \times \text{DAY}_{Sht},
\]

(2.4.16)

where \( \text{DAY}_{Sht} \) denotes the number of days that have passed since the last category purchase and \( INV_{h0} \) is set to zero. Let \( \gamma_h = (\gamma_{h1}, \ldots, \gamma_{h,3+2J})' \) be the vector of household specific coefficients to take into account individual heterogeneity and \( \epsilon_{ht} \) be the error term distributed as \( \mathcal{N}(0, 1) \).

We expect \( CR_h \) to positively influence the incidence decision. A household with a higher consumption rate would purchase the product more often to keep up with the pace of consumption. Thus, it is more likely that a category purchase occurs at each store visit. We expect \( INV \) to have a negative effect on incidence. When the inventory level is high, a further purchase would incur the costs of holding inventory. On the other hand, if there is not enough inventory left, there is a risk of stocking out, and the household would face pressure to replenish its inventory. We also expect that \( \text{PRICE}_{htj} \) will have a negative effect and that \( \text{FEATURE}_{htj} \) will have a positive effect.

### 2.4.2 Brand Choice

The random coefficient multinomial probit model is adopted for the brand choice decision. The probability that the household chooses brand \( j \) is given by

\[
B_{htj} = \text{Prob}(v_{htj} > v_{hti}, \forall i \neq j),
\]

(2.4.17)

where \( v_{htj} \) is the unobserved utility of household \( h \) for brand \( j \) at \( t \). We observe \( b_{ht} = j \) if and only if \( v_{htj} > \max_{i \neq j} v_{hti} \). We assume

\[
v_{htj} = \sum_{k=2}^{J} \delta_{h,k-1} D_k + \delta_{h,j} \text{PRICE}_{htj} + \delta_{h,J+1} \text{FEATURE}_{htj} + \eta_{htj},
\]

(2.4.18)

where \( D_k, k = 2, \ldots, J \) is the brand specific intercept, which equals 1 if \( k = j \) and \( D_1 = 0 \) for identification, \( \eta_{htj} \) is the error term associated with brand \( j \), and \( \delta_h = (\delta_{h1}, \ldots, \delta_{h,J+1})' \) is the vector of the household-specific coefficients. Finally, it is assumed that \( \eta_{ht} = (\eta_{ht1}, \ldots, \eta_{htJ})' \sim \mathcal{N}(0, \Omega) \) with \( \Omega = \text{diag}(1, \omega_2, \ldots, \omega_J) \). The diagonal covariance matrix is simple, but it is still able to avoid the independence of irrelevant alternatives. We do not use the differenced utilities since (2.4.18) is more
convenient for computing the choice probabilities that are required to evaluate the copula (Rossi et al., 1996). As in the incidence model, we expect that \( PRICE_{htj} \) negatively influences the brand choice decision and \( FEATURE_{htj} \) positively influences the brand choice decision.

Zhang et al. (2005) used the nested logit model for incidence and choice. However, Chib et al. (2004) argued that it can be restrictive. The nested logit model inserts the unobserved utilities for brand choice into the utility for incidence through the inclusive value measure. This implies that the marketing mix variables in the brand choice model have identical effects in the incidence model. In addition, the brand that has the greatest brand specific intercept influences the incidence decision most. Furthermore, as mentioned above, it is impossible to measure the dependence between incidence and choice; thus, the conditional probabilities cannot be obtained.

### 2.4.3 Quantity Choice

Finally, we specify the model for quantity choice. Following Bucklin et al. (1998) and Zhang et al. (2005), we use the random coefficient truncated Poisson regression model since it always produces positive and discrete outcomes. We assume that purchase quantity follows the zero truncated Poisson distribution, whose distribution function is given by

\[
Prob(q \leq q_{ht}) = \sum_{k=1}^{q_{ht}} \frac{\lambda_{ht}^k}{k!(\exp(\lambda_{ht}) - 1)}, \quad q_{ht} = 1, 2, \ldots
\]  

where \( \lambda_{ht} > 0 \) is the rate parameter. We further assume that \( \lambda_{ht} \) depends on the covariates as

\[
\lambda_{ht} = \exp(\theta_{h1} + \theta_{h2} CR_h + \theta_{h3} DEVINV_{ht} + \theta_{h4} PRICE_{htj} + \theta_{h5} FEATURE_{htj}),
\]

where \( DEVINV_{ht} \) is the deviation of the inventory level of household \( h \) at \( t \) from the average inventory level and \( \theta_h = (\theta_{h1}, \ldots, \theta_{h5})' \) is the vector of household-specific coefficients. We expect that \( CR_h \) and \( FEATURE_{htj} \) have positive effects and that \( PRICE_{htj} \) has a negative effect on quantity choice. Similar to \( INV_{ht} \) in the incidence model, the coefficient of \( DEVINV_{ht} \) is expected to be negative.

### 2.5 Estimation

We assume that the household specific coefficients \( \beta_h = (\gamma_h', \delta_h', \theta_h')' \) follow \( N(\beta, B) \), where \( \beta \) and \( B \) are the unknown parameters. The structure of \( B \) could be of interest, since this matrix gives the covariance between the coefficient heterogeneity in the incidence, choice, and quantity models (Chib et al., 2004).

Following Zhang et al. (2005), the likelihood is given by

\[
L = \prod_{h=1}^{H} \prod_{t=1}^{T_h} \left[ (P_{htj})^{y_{ht}} (1 - Y_{ht})^{1 - y_{ht}} \right].
\]
The first term in the brackets corresponds to the joint probability of the purchase outcome and the second term is the probability of non-purchase. The parameters are estimated using the Markov chain Monte Carlo method (MCMC) in the Bayesian framework (see e.g., Rossi et al., 2006). In Appendix A, the prior distribution for the parameters are specified and the resulting MCMC method is described.

2.6 Elasticity

We show how the conditional probabilities derived in Section 2 are utilized to decompose the overall price elasticity in our model framework.

Zhang et al. (2005) derived the elasticities based only on the marginal models. The elasticities for incidence, brand choice, and quantity are given by

\[
\frac{d \text{Prob}(y_{ht} = 1)}{d \text{PRICE}_{htj}} \cdot \frac{\text{PRICE}_{htj}}{\text{Prob}(y_{ht} = 1)},
\]

(2.6.21)

\[
\frac{d \text{Prob}(b_{ht} = j)}{d \text{PRICE}_{htj}} \cdot \frac{\text{PRICE}_{htj}}{\text{Prob}(b_{ht} = j)}, \quad \text{and} \quad \frac{d \text{Prob}(q = q_{ht})}{d \text{PRICE}_{htj}} \cdot \frac{\text{PRICE}_{htj}}{\text{Prob}(q = q_{ht})},
\]

respectively. Since these elasticities are not derived following the sequence of purchase decisions, they can be infeasible.

In contrast, we can fully exploit the structure of the Gaussian copula. We are interested in how the household reacts to the price change at each stage of the sequence of decisions. Specifically, it is necessary to derive the following price elasticities for brand choice and quantity, which are given by

\[
\frac{d \text{Prob}(b_{ht} = j | y_{ht} = 1)}{d \text{PRICE}_{htj}} \cdot \frac{\text{PRICE}_{htj}}{\text{Prob}(b_{ht} = j | y_{ht} = 1)},
\]

(2.6.22)

and

\[
\frac{d \text{Prob}(q = q_{ht} | y_{ht} = 1, b_{ht} = j)}{d \text{PRICE}_{htj}} \cdot \frac{\text{PRICE}_{htj}}{\text{Prob}(q = q_{ht} | y_{ht} = 1, b_{ht} = j)},
\]

(2.6.23)

respectively. We use the same incidence elasticity as in (2.6.21). The elasticities in (2.6.22) and (2.6.23) can be calculated using the conditional probabilities in (2.3.11) and (2.3.12), respectively. In Appendix B, we detail the derivation of the elasticities based on our marginal models specified in Section 2.4. The elasticities are estimated using the MCMC output.

Our approach is particularly interesting because the elasticity compute from (2.6.22) reveals the impact of price changes on brand choice given that the incidence occurs, while the elasticity computed form (2.6.23) reveals the impact of price changes on quantity choice given the incidence and choice decisions. Thus, manufacturers and retailers may increase their profits by promotions that target a particular step in the decision process. For example, if the quantity choice decision is more sensitive to the price conditionally on the incidence and choice decisions than the incidence decision, a retailer would expand its sales by trying to induce households to purchase more units...
Table 1: Summary of the response variables

<table>
<thead>
<tr>
<th></th>
<th>Brand 1</th>
<th>Brand 2</th>
<th>Brand 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of purchases from the category</td>
<td>4814</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand Share</td>
<td>0.390</td>
<td>0.252</td>
<td>0.358</td>
</tr>
<tr>
<td>Quantity Mean</td>
<td>1.257</td>
<td>1.487</td>
<td>1.416</td>
</tr>
<tr>
<td>SD</td>
<td>0.530</td>
<td>0.634</td>
<td>0.656</td>
</tr>
<tr>
<td>Max</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

at each shopping occasion. If households respond to the price change by switching brands given the incidence decision, a manufacturer would need to focus on preventing its customers from switching to another brand and on trying to steal shares from competing brands.

2.7 Prediction

Utilizing the decomposition, we can predict the future outcomes given the data from the predictive distribution

\[
\text{Prob}(y_{ht^*} = 1, b_{ht^*} = j, q \leq q_{ht^*}|\text{Data}),
\]

for household \( h \) at time \( t^* > T_h \). In Appendix C, we show the MCMC method for sampling from the predictive distribution.

3. Empirical Analysis

3.1 Data Description

We apply the proposed model to the scanner panel data provided by CUSTOMER COMMUNICATIONS, Ltd. The data include 22,365 observations for 594 households that purchased from the dishwashing category at a drugstore in Japan between 1 January 2002 and 30 June 2003. The number of observations per household ranges between 31 and 95. The product category consists of three brands. Since the brand names are not available to us, we denote them as Brand 1, Brand 2, and Brand 3. For each household, we randomly split the data in such a way that 70% of the observations are included in the calibration data for calibrating the model, while the outcomes of the remaining 30% are to be predicted and kept in the holdout data. The calibration data and holdout data consist of 15,383 and 6,982 observations, respectively. The summary of the response variables and explanatory variables in the calibration data are given by Tables 1 and 2, respectively. It is observed from Table 1 that category purchases occurred in very small quantities, as the maximum number of purchase quantity is six for all the brands. Thus, it would be reasonable to use the Poisson regression model for quantity choice.
Table 2: Summary of the explanatory variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CR_h$</td>
<td>0.051</td>
<td>0.034</td>
</tr>
<tr>
<td>$INV_{ht}$</td>
<td>0.232</td>
<td>2.264</td>
</tr>
<tr>
<td>$DEV_{INV_{ht}}$</td>
<td>-0.015</td>
<td>1.468</td>
</tr>
<tr>
<td>$PRICE_{ht1}$</td>
<td>0.454</td>
<td>0.037</td>
</tr>
<tr>
<td>$PRICE_{ht2}$</td>
<td>0.439</td>
<td>0.061</td>
</tr>
<tr>
<td>$PRICE_{ht3}$</td>
<td>0.403</td>
<td>0.048</td>
</tr>
<tr>
<td>$FEATURE_{ht1}$</td>
<td>0.027</td>
<td>0.163</td>
</tr>
<tr>
<td>$FEATURE_{ht2}$</td>
<td>0.071</td>
<td>0.257</td>
</tr>
<tr>
<td>$FEATURE_{ht3}$</td>
<td>0.069</td>
<td>0.253</td>
</tr>
</tbody>
</table>

3.2 Model Validity

We computed the log marginal likelihood to compare the model fit to the data. As alternative models, we consider the restricted version of the proposed model with the dependence matrix given by (2.3.13), which corresponds to the setting of Zhang et al. (2005), and the unified model of Song and Chintagunta (2006) as considered in Ebling and Klapper (2010). The log marginal likelihoods are $-11394.072$, $-19463.295$, and $-13901.936$ for the proposed model, restricted model, and unified model, respectively.

As discussed in Section 3.3.1, suppressing $\rho_{12}$ to zero causes substantial deterioration of the model fit, and thus, we use the general dependence structure for the present dataset. Since the unified model has a single utility function to be maximized, it can be very restrictive when households respond to the marketing mix in different ways across the three decisions, and the model does not guarantee that the observed outcomes provide the highest utility (see Ebling and Klapper, 2010). Therefore, the result that the proposed model is superior to the unified model implies that the responsiveness of the households to the marketing mix in the dishwashing category varies considerably across the three decisions, as discussed below.

It may be argued that we need to take the segmentation of households into account since the households in one segment may reveal a different purchase pattern and an interdependence pattern from households in another segment, as in Bucklin et al. (1998). However, since dishwashing soaps are necessity goods, there would not be much difference in purchase behavior between the households. Therefore, introducing random coefficients is sufficient to absorb the individual heterogeneity. We also calculated the log marginal likelihood of the model with two segments based on the finite mixture approach. The log marginal likelihood for the model with two segments is $-12801.718$. The single segment model is decisively favored by the data over the model with two segments.

The predictive performance of the proposed model is compared with the unified model. It is measured by the fraction of correct predictions of the responses in the holdout data. The fractions of correct predictions for the proposed and the unified models are given in Table 3. The proposed model has some comparable predictive
performance to the unified model. The proposed model outperformed the unified model for the cases of incidence and brand choice, but it performed slightly worse than the unified model for the case of quantity choice.

3.3 Parameter Estimates

3.3.1 Copula Parameters

The posterior means, standard deviations, and 95% credible intervals for the components of $\mathbf{R}$ are presented in Table 4. The posterior mean of $\rho_{13}$ is $-0.559$, which implies a relatively strong negative dependence between the incidence and quantity choice decisions. Thus, purchases from the dishwashing soap category are associated with smaller purchase quantities. This would imply that purchases from this category do not induce stockpiling and that households buy the products simply to replenish their inventories. This result is in line with the observation that households buy dishwashing soaps only in very small purchase quantities. This negative incidence–quantity dependence was also observed for the paper towel category in Zhang et al. (2005).

Similarly, the posterior means of $\rho_{23}$ are $-0.553$, $-0.529$, and $-0.551$ for Brands 1, 2, and 3, respectively. We thus found negative dependences between the brand choice and quantity choice decisions for all three brands because there are very strong positive dependences between the incidence and brand choice decisions, as discussed below.

The posterior means of $\rho_{12}$ resulted in similar values, which are near the upper boundary of the supports. This result suggests that the choice of each brand is

---

1) We would like to leave the further analysis of the predictive performance for future study.

---

Table 3: Fractions of correct predictions

<table>
<thead>
<tr>
<th>Incidence</th>
<th>Proposed</th>
<th>Unified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand 1</td>
<td>0.703</td>
<td>0.313</td>
</tr>
<tr>
<td>Brand 2</td>
<td>0.608</td>
<td>0.400</td>
</tr>
<tr>
<td>Brand 3</td>
<td>0.628</td>
<td>0.440</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Proposed</th>
<th>Unified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.658</td>
<td>0.705</td>
</tr>
</tbody>
</table>

Table 4: Posterior summary for $\mathbf{R}$

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{13}$</td>
<td>$-0.559$</td>
<td>0.001</td>
<td>($-0.560$, $-0.557$)</td>
</tr>
<tr>
<td>Brand 1</td>
<td>$\rho_{12}$</td>
<td>0.998</td>
<td>(0.997, 0.998)</td>
</tr>
<tr>
<td></td>
<td>$\rho_{23}$</td>
<td>$-0.553$</td>
<td>($-0.554$, $-0.553$)</td>
</tr>
<tr>
<td>Brand 2</td>
<td>$\rho_{12}$</td>
<td>0.990</td>
<td>(0.985, 0.993)</td>
</tr>
<tr>
<td></td>
<td>$\rho_{23}$</td>
<td>$-0.529$</td>
<td>($-0.534$, $-0.514$)</td>
</tr>
<tr>
<td>Brand 3</td>
<td>$\rho_{12}$</td>
<td>0.997</td>
<td>(0.996, 0.997)</td>
</tr>
<tr>
<td></td>
<td>$\rho_{23}$</td>
<td>$-0.551$</td>
<td>($-0.552$, $-0.549$)</td>
</tr>
</tbody>
</table>
strongly tied to the incidence decision. This may reflect the strong favoritism of households toward the brands. Although the increases in the brand choice probability and incidence probability are both associated with smaller purchase quantities, the increase in the brand choice probability is strongly and positively tied to the increase in the incidence probability. Therefore, this drugstore would benefit from promoting any of the brands and trying to induce households to purchase from the category at each store visit, but not from trying to induce households to buy more units. From the viewpoint of the manufacturers of the category brands, since incidence choice is strongly tied to brand choice, a brand may steal shares from competing brands by implementing a sales promotion that specifically targets the incidence decision of households.

### 3.3.2 Marginal Models

Table 5 shows the posterior means, standard deviations, and 95% credible intervals for the coefficients $\gamma$ in the incidence model. The table shows that the posterior means of all the coefficients have the expected signs. $CR_h$ and $INV_{ht}$ have credible positive and negative effects, respectively, on the incidence decision: their posterior means are 3.926 and $-0.147$, respectively, and the credible intervals do not include zero. Thus, it is implied that households whose consumption rates are high and whose inventory levels are low are more likely to buy from the category at a given shopping occasion. For all three brands, it is found that $PRICE_{htj}$ credibly and negatively influences the incidence decision. The posterior means of Brands 1, 2, and 3 are $-0.825$, $-0.307$, and $-0.508$, respectively. The magnitudes of the coefficients for the prices are the largest for Brand 1 and the smallest for Brand 2. We will investigate price elasticity in Section 2.6. The posterior means of the coefficients for $FEATURE_{htj}$ are all positive and the credible intervals do not include zero for all brands. Therefore, when some brands of the category are promoted at a shopping occasion, households are more likely to purchase from the category.

Table 6 shows the posterior means, standard deviations, and 95% credible intervals for the coefficients $\delta$ in the brand choice model. The posterior mean of $D_2$ is $-0.608$ and that of $D_3$ is 0.307. Therefore, the baseline preference for Brand 2 is the lowest and that for Brand 3 is the highest of the three brands. The posterior
means for \( PRICE_{htj} \) and \( FEATURE_{htj} \) are \(-3.027\) and \(2.364\), respectively, with the expected signs and the credible intervals do not include zero. Therefore, the price and promotion of the brand have negative and positive effects, respectively, on the brand choice.

Table 7 shows the posterior means, standard deviations, and 95% credible intervals for the coefficients \( \theta \) in the quantity choice model. The results given in the table show the expected signs. The posterior means of \( CR_h \) and \( DEVINV_{ht} \) are \(4.777\) and \(-0.502\), respectively. Households are more likely to buy in larger quantities when their speed of consumption is fast and when the inventory level is below average. The posterior mean of \( PRICE_{htj} \) is \(-5.944\), and the credible interval does not include zero. The magnitude of the coefficient seems quite large, which may suggest that households are very sensitive to the price change in the purchase quantity choice. However, in Section 2.6 it is demonstrated that this is not the case when conditioned on the incidence and choice decisions. Finally, the table shows that the posterior mean of \( FEATURE_{htj} \) is negative and that the credible interval does not include zero. A possible explanation for this result is as follows. When households face a promotion in the product category at a shopping occasion, they infer either that the sale price will reappear soon again or that the sale price will not reappear until a certain period of time has passed. While the latter inference would induce households to buy in larger quantities and to form stockpile until the next promotion, the former would discourage them to do so since the attractiveness of the current promotion is decreased (Zhang et al., 2012). Therefore, the result implies that while the promotion would encourage households to purchase from the category and choose the brand on sale, they tend to buy the products in smaller quantities because they expect the future price to be low.

### 3.4 Elasticity Analysis

Table 8 shows the posterior medians of the price elasticities for the three brands.
Table 8: Price elasticities

<table>
<thead>
<tr>
<th>Proposed</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence</td>
<td>Quantity</td>
</tr>
<tr>
<td>Brand 1</td>
<td>−0.491</td>
</tr>
<tr>
<td>Brand 2</td>
<td>−0.221</td>
</tr>
<tr>
<td>Brand 3</td>
<td>−0.315</td>
</tr>
</tbody>
</table>

The figures shown in the second and third columns are computed using the conditional probabilities based on (2.6.22) and (2.6.23), respectively, and the elasticities in the fourth and fifth columns are computed based on the marginal probabilities in (2.6.21).

For the case of the incidence decision, the price elasticities with respect to the incidence probability are −0.491, −0.221, and −0.315 for Brands 1, 2, and 3, respectively. The incidence decision is the most sensitive to the price change in Brand 1 and is least sensitive to the price change in Brand 2, as suggested by the magnitudes of the corresponding coefficients for $PRICE_{htj}$.

While the conditional and marginal price elasticities for Brand 1 resulted in similar values, those for Brand 2 are −0.373 and −0.287 and those for Brand 3 are −0.171 and −0.237, respectively. The differences between the conditional and marginal elasticities are 0.086 for Brand 2 and 0.066 for Brand 3. These are notable differences and hence, using marginal elasticity instead of conditional elasticity would result in an overstated or understated conclusion on elasticity. In addition, the magnitude of conditional elasticity is the largest for Brand 2. Since the share for Brand 2 is the smallest of the three, this result supports the finding in the literature that the demand for brands with smaller shares is more sensitive to the price changes than that for brands with larger shares.

There are substantial differences in the estimated quantity elasticities. The elasticities computed based on the conditional probabilities are much smaller in magnitude than those based on the marginal probabilities, and they are close to zero. The marginal elasticities for quantity choice are −0.479, −0.816, and −0.593 for Brands 1, 2, and 3, respectively. This may suggest that households are very sensitive to price changes when they determine their purchase quantity based on these marginal elasticities. However, the conditional elasticities for quantity choice are minuscule: 0.003, −0.019, and −0.003 for Brands 1, 2, and 3, respectively. This result suggests that households do not respond to price changes when they choose their quantity after deciding to buy from the category and which brand to buy. This is a reasonable result since the most of the category purchases occurred in small quantities. Furthermore, since dishwashing soaps are necessity goods, when households are low on inventory, they have to buy at least one unit of the product of the chosen brand. Therefore, the main sources of the overall behavioral response to the price change are the incidence and brand choice decisions, but the responsiveness of the quantity decision to the price change hardly contributes to the overall response. Consequently, retailers and manufacturers should not expect price promotion to lead to increases in purchase
quantity. This result also conforms to the argument in Section 3.3.1.

4. Conclusion

We proposed an integrated purchase model of category purchase incidence, brand choice, and quantity choice using the Gaussian copula. Our parameterization of the Gaussian copula enabled us to decompose the joint probability into a sequence of conditional probabilities that can be used to compute price elasticities and to predict future outcomes that reflect the interdependence among the decisions.

Using the scanner panel data for the dishwashing soap category, we found that there exist strong interdependences among incidence, brand choice and quantity decisions. The incidence and brand choice decisions were found to be the main sources of the overall elasticity, while quantity choice was hardly elastic to the price change conditional on the incidence and choice decisions.

Possible directions for future research are as follows. First, since our results suggest that there exists strong interdependence among the decisions, especially between the incidence and brand choice decisions, it would be of interest to study what drives interdependence. Second, in this paper we decomposed the joint probability into conditional probabilities in the order of incidence, brand choice, and quantity. Since decomposition into another order may reflect the sequence of the purchase decisions more appropriately, it would be interesting to compare the orders in terms of, for example, predictive performances. Finally, if we are interested in the timing by which the incidence occurs, it is possible to incorporate a purchase timing model in place of the binary probit model used in this paper. Furthermore, there exists an extensive body of literature on multiple category purchase models. It is also possible to extend our copula modeling approach to the multiple category case.

Acknowledgements

The authors thank Takuya Satomura (Keio University) his the valuable comments at the 90th Japan Institute of Marketing Science Conference. They also thank the editor and two referees for improving the present manuscript. This research is partly supported by the Grants-in-Aid for JSPS Fellows (DC1) and the Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology (#25245035). The computational results are obtained using Ox version 6.21 (Doornik, 2007).

REFERENCES


Appenendix A: Estimation Procedure

To complete the model formulation, we specify the prior distributions for unknown parameters. Particularly, we assume

\[ \beta \sim N(\beta_0, \Sigma_0), \]
\[ B \sim IW(\nu_0, B_0), \]
\[ \rho_{12}^j, \rho_{13}^j, \rho_{23}^j \sim TN(-1,1)(0,1), \quad j = 1, \ldots, J, \]
\[ \omega_{jj} \sim IG(n_0, s_0), \quad j = 2, \ldots, J. \]

where \( IW(\nu, \Psi) \) denotes the inverse Wishart distribution with \( \nu \) degrees of freedom and the scale matrix \( \Psi \), \( TN(a,b)(\mu, \sigma^2) \) denotes the normal distribution with mean \( \mu \) and variance \( \sigma^2 \) truncated on the interval \( (a, b) \), and \( IG(\kappa, \zeta) \) denotes the inverse gamma distribution with parameters \( \kappa \) and \( \zeta \). The values of the hyperparameters are chose so that the prior distributions are relatively flat.

Our MCMC scheme consists of the random walk Metropolis–Hastings (MH) algorithm and the univariate slice sampler (Neal, 2003). Specifically, we sample the parameter values in the following way:

\( \beta_h : \) the MH algorithm is applied for \( h = 1, \ldots, H : \)
\( \beta : \) sampled from \( N(\hat{\beta}, \hat{\Sigma}) \),
\[ \text{where } \hat{\Sigma} = \left( HB^{-1} + (\Sigma)^{-1} \right)^{-1} \text{ and } \hat{\beta} = \hat{\Sigma} \left( B^{-1} \sum_{h=1}^{H} \beta_h + \Sigma_0^{-1} \beta_0 \right) : \]
\( B : \) sampled from \( IW(\hat{\nu}, \hat{B}) \),
\[ \text{where } \hat{\nu} = \sum_{h=1}^{H} T_h + \nu_0 \text{ and } \hat{B} = \sum_{h=1}^{H} (\beta_h - \beta)(\beta_h - \beta)^t + B_0 : \]
\( \rho_{12}, \rho_{13}, \rho_{23} : \) slice sampling is applied for \( j = 1, \ldots, J : \)
\( \omega_{jj} : \) the MH algorithm is applied for \( j = 2, \ldots, J : \)

In our analysis, the hyperparameters are set to the following values: \( \beta_0 = 0, \Sigma_0 = 100I, \nu_0 = 3J + 9 + 1 = 19, B_0 = 0.1I, n_0 = 3, s_0 = 1. \) We ran MCMC with some reasonable starting values for 40,000 iterations with an initial burn-in of 20,000
iterations.

Appendix B: Elasticity

We derive the price elasticities in (2.6.21), (2.6.22), and (2.6.23) based on the marginal models specified in Section 2.4. For simplicity, we rewrite the conditional probability in (2.3.11) and (2.3.12) by

\[ \Phi_{21}(\Phi^{-1}(B_{htj})) \quad \text{and} \quad \Phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht}))) - \Phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht} - 1))), \]

where \( \Phi_{21}(\cdot) \) and \( \Phi_{312}(\cdot) \) are the distribution functions of \( \mathcal{N}(\rho_{12}^1 \Phi^{-1}(Y_{ht})1 - (\rho_{12}^1)^2) \) and \( \mathcal{N}(\mathbf{r}_{21j} \mathbf{R}_{11j}^{-1} \mathbf{x}_j, 1 - \mathbf{r}_{21j} \mathbf{R}_{11j}^{-1} \mathbf{r}_{12j}) \) with \( \mathbf{x}_j = (\Phi^{-1}(Y_{ht}, \Phi^{-1}(B_{htj})) \)' and whose density functions are denoted by \( \phi_{21}(\cdot) \) and \( \phi_{312}(\cdot) \), respectively.

First, the price elasticity for incidence based on our binary probit model is given by

\[ \gamma_{ht,3+j} \phi(z_{ht}) \frac{\text{PRICE}_{htj}}{\Phi(z_{ht})}, \quad (B1) \]

where \( \phi \) is the density function of \( \mathcal{N}(0,1) \). Second, we show the derivative of the conditional choice probability with respect to \( \text{PRICE}_{htj} \). This is given by

\[ \frac{d\text{Prob}(b_{ht} = j | y_{ht} = 1)}{d\text{PRICE}_{htj}} = \frac{d\Phi_{21}(\Phi^{-1}(B_{htj}))}{d\text{PRICE}_{htj}} \]

\[ = \frac{d\Phi_{21}(\Phi^{-1}(B_{htj}))}{d\Phi^{-1}(B_{htj})} \frac{d\Phi^{-1}(B_{htj})}{d\text{PRICE}_{htj}} \frac{dB_{htj}}{d\text{PRICE}_{htj}} \]

\[ = \frac{\phi_{21}(\Phi^{-1}(B_{htj}))}{\phi(\Phi^{-1}(B_{htj}))} \frac{dB_{htj}}{d\text{PRICE}_{htj}}. \quad (B2) \]

The second term in (B2) can be computed using the GHK simulator (Kim et al., 2003). Finally, the derivative of the conditional probability of quantity with respect to \( \text{PRICE}_{htj} \) is given by

\[ \frac{d\text{Prob}(q_{ht} = q | y_{ht} = 1, b_{ht} = j)}{d\text{PRICE}_{htj}} = \frac{d\Phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht})))}{d\text{PRICE}_{htj}} \frac{d\Phi^{-1}(\text{Prob}(q \leq q_{ht}))}{d\text{Prob}(q \leq q_{ht})} \]

\[ - \frac{d\Phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht} - 1)))}{d\text{PRICE}_{htj}} \frac{d\Phi^{-1}(\text{Prob}(q \leq q_{ht} - 1))}{d\text{Prob}(q \leq q_{ht})}. \quad (B3) \]

We only show the calculation of the first term on the right hand side of (B3): the second term can be calculated in a similar manner. It is given by

\[ \frac{d\Phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht})))}{d\text{PRICE}_{htj}} \]

\[ = \frac{d\Phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht})))}{d\Phi^{-1}(\text{Prob}(q \leq q_{ht}))} \frac{d\Phi^{-1}(\text{Prob}(q \leq q_{ht}))}{d\text{Prob}(q \leq q_{ht})} \frac{d\text{Prob}(q \leq q_{ht})}{d\text{PRICE}_{htj}} \]

\[ = \frac{\phi_{312}(\Phi^{-1}(\text{Prob}(q \leq q_{ht})))}{\phi(\Phi^{-1}(\text{Prob}(q \leq q_{ht})))} \frac{d\text{Prob}(q \leq q_{ht})}{d\text{PRICE}_{htj}}. \quad (B4) \]
where the second term of the product is given by

$$\sum_{k=1}^{\eta_h} \left( \frac{k^k \lambda^k (\exp(\lambda_{ht}) - 1) - \lambda^{k+1} \exp(\lambda_{ht})}{k! (\exp(\lambda_{ht}) - 1)^2} \right) \theta_{ht},$$

Combining the results of (B2), (B3), and (B4) in (2.6.22) and (2.6.23), the elasticities can be estimated using the MCMC method described in Appendix A.

**Appendix C: Prediction**

The probability in (2.7.24) can be written by

$$\text{Prob}(y_{ht*} = 1, b_{ht*} = j, q \leq q_{ht*} | \text{Data}) = \int \text{Prob}(y_{ht*} = 1) \text{Prob}(b_{ht*} = j | y_{ht*} = 1) \text{Prob}(q \leq q_{ht*} | y_{ht*} = 1, b_{ht*} = j) \pi(\Xi | \text{Data}) d\Xi,$$

$$= \int \Phi(Y_{ht*}) \Phi_2^{-1}(B_{ht*,j}) \Phi_3^{-1}(\text{Prob}(q \leq q_{ht*})) \pi(\Xi | \text{Data}) d\Xi, \quad (C5)$$

where $\Xi$ contains all the parameters and $\pi(\Xi | \text{Data})$ is the posterior distribution of $\Xi$. MCMC is implemented to draw from the predictive distribution (C5). The future quantity can be simulated by inverting the conditional probability.

(Received February 17 2014, Revised May 7 2014)