DISCUSSION PAPER SERIES
June 2014

No.2014-07

The Consideration Set Based Submarket Identification: An Application of the Hierarchical Factor Analysis Model

Sotaro Katsumata
Akihiro Nishimoto

The Consideration Set Based Submarket Identification: An Application of the Hierarchical Factor Analysis Model

Sotaro Katsumata
(Nagasaki University)
Akihiro Nishimoto
(Kwansei Gakuin University)

Abstract
In this study, we focus on the consumers' consideration set and submarkets of the market. In the modern matured market, there are too many competitive brands, and consumers cannot evaluate all the available brands. They only see the subset of the brands to compare them and choose between them. The subset of comparison and evaluation is called a consideration set. Brands are required to be in this consideration set, to be chosen. Therefore, identification of the consideration set is an important issue for brand managers. However, examining the consideration set is relatively difficult as compared with actual choice, because we cannot observe the composition of the set from behavioral data such as purchase records. To address this issue, we examine the consideration set collected through a research survey. In our analysis, we propose a model to examine the submarket structure. Our model incorporates factor analysis into a discrete choice model and assumes hierarchical structure. To use the discrete choice model, we can obtain the consideration probability of each brand. Factor analysis model enables us to identify submarkets. Since we assume the hierarchical expression, we can examine the relationship between consumers' characteristics and the consideration probability of each submarket.

Keywords
Consideration Set; Submarket; Discrete Choice Model; Factor Analysis; Hierarchical Model

1. Introduction

Consumers' choice is one of the most important issues in marketing and consumer research. One of the goals of every firm is to be the preferred choice of consumers. If the firm products were not chosen, the firm would not grow. However, the choice behavior of real consumers undergoes a complex process. The complexity occurs because of the selection of alternatives. Many previous studies have indicated the existence of the choice subset. For example, Simon (1947) defined the behavioral model of a human based on the cognitive limitation. According to Simon (1947), humans evaluate only a few alternatives rationally and choose a "satisfied" alternative in the subset even if there are more preferable alternatives out of the subset. Further, Howard and Sheth (1969) mention the existence of the "evoked set" as a kind of choice subset within their consumer behavioral model. Firm managers have to consider the existence of the choice subset and recognize that it is the first step in the process of consumers choosing their brand. The consideration set is especially regarded as one of the most important subsets. As many previous studies such as Shocker et al. (1991), Roberts and Lattin (1991, 1997), and Elrod and Keane (1995) have highlighted, one of the most important choice subsets is the consideration set. Consumers carefully evaluate the brands within the consideration set. Therefore, being in the consideration set of consumers is an important milestone for managers. Thus, in this study, we focus on the consideration set.

Consumers tend to evaluate and compare rationally; the choice problem within the consideration set is relatively simple and consistent. However, it is difficult to identify which alternatives are in the subset. Moreover, the number of candidate brands becomes large in the modern market. One of the clues to identifying the alternatives in the subset is the submarket structure. There are some submarkets and brands within the same submarket, which tend to be compared to each other by consumers.

In this study, we focus on the consideration set and the market structure using the stated consideration data. We collected the consideration dataset as a multivariate binomial observation from research survey and we construct a model to explain it. We apply the factor analysis model to express the submarket of the market. Additionally, we assume the hierarchical structure to examine the relationship between each factor (or consumer characteristic) and consideration set memberships.

2. Previous Studies on Choice Set

In this section, we briefly survey the previous studies and models of the consideration set. Let $P_i(j)$ be a choice probability of brand j by consumer i and let his/her choice set be C_i. The general formulation is given by Manski (1977) and Shocker et al. (1991) as follows:

$$P_i(j) = \sum_{C \in G(j)} P_i(j|C) P_i(C|G) \tag{1}$$

where G is the set of all the possible consideration sets and $G(j)$ is the set of all elements of G that contain brand j. This expression suggests that the choice probability of brand j consists of two conditional probabilities. $P_i(C|G)$ is the probability that the choice set C is formulated by consumer i, and $P_i(j|C)$ is the probability that brand j is chosen from among the given choice set C. From this expression, we can estimate the choice probabilities for each customer i.

For example, if there is a set of available brands $\{a, b, c\}$, all of the possible choice sets are $\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}$, and $\{a, b, c\}$ (except the empty set). The set of choice sets G consists of these seven subsets. For brand a to be chosen, the choice set needs to be $\{a\}, \{a, b\}, \{a, c\}$, or $\{a, b, c\}$. These four subsets are elements of $G(a)$. Based on this formulation, some empirical models are provided, such as those of Andrews and Srinivasan (1995), Siddarth et al. (1995), and Chiang et al. (1999). However, this formulation has a serious limitation. The number of possible subsets becomes very large as the number of alternative brands increases. When the number of alternatives is J, the number of choice subset is $2^J - 1$. Therefore, we have to estimate probabilities for $2^J - 1$ subsets.

To address this problem, Nierop et al. (2010) propose a model that estimates the probability of a choice subset considered by a consumer for each alternative. In this model, the membership probability of the choice subset is estimated using a binomial discrete choice model. Therefore, this model estimates the membership probabilities from their latent variables. In this probabilistic choice model, if the number of available brands is J, the number of the probability is J. However, this model has a possibility that the empty set occurs. Nierop et al. (2010) introduce an assumption that the brand that has highest membership probability shall automatically be an element of the choice set even if the latent variable of the membership probability does not exceed a certain threshold.

In this study, we also assume the membership probability of the choice set based on Nierop et al. (2010). Thus, we suppose that the observed choice probability of brand j by consumer i is obtained from the product of two probabilities. The first term, $P_i(j \in C_i)$, is the probability of j being an element of the choice set. The second term, $P_i(j|C_i)$, is the probability of j, chosen from among the given choice set C_i.

$$P_i(j, j \in C_i) = P_i(j|C_i) \times P_i(j \in C_i) \tag{2}$$

As mentioned in the previous section, our research purpose is to examine the formulation of the choice set. Therefore, in this study, we only focus on the right term, $P_i(j \in C_i)$. Of course, we adopt our model to the conditional choice model such as the incidence, choice, and quantity model proposed in Bucklin et al. (1998) and van Heerde and Neslin (2008). In the following section, we propose a model based on the factor analysis model.

3. The Probabilistic Model of the Consideration Set

3.1. Consideration Probability

At first, we define an observed indicator variable z_{ij}, where brand j is an element of C_i, which is the consideration set of consumer i.

$$z_{ij} = \begin{cases} 1 & if \quad j \in C_i \\ 0 & else \end{cases} \tag{3}$$

The observation variable z_{ij} is a binary measure so that if consumer i has brand j in his/her choice set, $z_{ij} = 1$; otherwise, 0. We assume the latent variable z_{ij}^* based on Albert and Chib (1993) as follows:

$$z_{ij} = \begin{cases} 1 & if \quad z_{ij}^* > 0 \\ 0 & else \end{cases} \quad (4)$$

To estimate the membership probability of each brand, we introduce the factor analysis model. The factor model has good properties to examine the submarket structure and consideration set composition (Roberts and Lattin 1991, 1997). For example, Elrod and Keane (1995) apply a discrete choice factor model to identify the submarket and draw a product map. Although we can draw a scatter plot of the market if the dimension of the factor is two or three, we do not aim to obtain an output figure in this study.

The factor model to explain the latent variable is as follows:

$$z_i^* = \alpha + \Lambda f_i + \varepsilon_i, \varepsilon_i \sim N(0, I_J) \quad (5)$$

where $z_i^* = \left(z_{i1}^*, \cdots, z_{iJ}^*\right)'$, Λ is a $J \times Q$ matrix parameter of factor loadings, and f_i is a $Q \times 1$ vector of factor scores. This is a Q-dimensional factor model. Since the objective variable is not standardized, we need to incorporate an intercept parameter α. In addition, the usual factor model requires the diagonal unknown variance parameters. However, to satisfy the identification condition, we assume an identity matrix for the variance-covariance matrix. From this model, we can estimate the consumer i's membership probability of brand j from the factor loading Λ and the factor score f_i.

The dimension Q is interpreted as the number of the submarkets of the objective competitive market. For example, if the value of λ_{jp}, the (j, p) element of the factor loadings parameter Λ, is high, brand j tends to be a member of submarket p. Note that we do not assume the deterministic submarket membership. Our assumption is a fuzzy submarket membership such as product mapping (Elrod and Keane 1995). This suggests that brands that have high values at the p-th column of Λ are under competition within the submarket p.

As regards consumers, we can examine the submarket level consideration (preference) by f_i. If the element f_{ip} is high, consumer i tends to consider brands that have a high value at p-th column of Λ. Since the factor score f_i is obtained for each consumer, it is a source of differentiation of the consideration probability of brand j. The latent variable of the membership probability of consumer i for brand j is the addition of the factor loadings and scores, as follows:

$$z_{ij}^* = \alpha_j + \lambda_{j1} f_{i1} + \cdots + \lambda_{jq} f_{iq} + \cdots + \lambda_{jQ} f_{iQ} + \varepsilon_{ij} \quad (6)$$

From the latent variable z_{ij}^*, we can estimate the probability that consumer i considers brand j in a purchase occasion, where $\phi(x|m, v^2)$ is a density function of the univariate normal distribution whose mean is m and variable is v^2.

$$P_i(j \in C_i) = \int_0^\infty \phi(x|\alpha_j + \lambda_j' f_i, 1)\, dx \quad (7)$$

In this study, we assume the prior structure for the model. Therefore, we relax the assumption of the factor scores, which is usually assumed as $f_i \sim N_Q(0, I_Q)$. In this study, we do not assume the mean as 0, and set the following linear regression model:

$$f_i = \Delta w_i + \zeta_i, \zeta_i \sim N_Q(0, I_Q) \tag{8}$$

where w_i is a consumer characteristics variable such as age or gender. In this model, these consumer characteristics affect the factor scores as a prior. In this study, we incorporate the consumer demographic variable and knowledge. The detailed setting is found in the next section.

Our factor analysis model, which assumes a discrete objective variable and hierarchical structure, has to adopt the Markov Chain Monte-Carlo (MCMC) method to estimate parameters. The detailed procedure of the estimation is in Appendix A.1.

3.2. Identification of the Factor Model

To obtain the estimates by the MCMC method, we have to satisfy the identification condition of the factor analysis. For example, Geweke and Zhou (1996) propose a way to satisfy the condition introducing the restriction in the factor loadings parameter. This restriction is also adopted by Lopes and West (2004). Using another approach, Ansari and Jedidi (2000) restrict the covariance matrix. In addition, this study incorporates a discrete objective variable. However, to estimate the restricted covariance matrix, we have to use the Metropolis-Hastings (M-H) method, and the computational load is relatively high. Therefore, we use the restriction proposed by Geweke and Zhou (1996).

In the model of Geweke and Zhou (1996), the factor loadings matrix is restricted to identify the factor loadings and scores, which are able to rotate. In other words, for any $J \times J$ matrix R, where $RR' = I$, $\Delta^* = P'\Delta$ and $f^* = P'f_i$ are also valid parameters. Therefore, to address the problem, we restrict Λ_Q, which is a $Q \times Q$ sub-matrix of $J \times Q$ factor loadings Λ and $\Lambda_Q = \Lambda_{1:Q,1:Q}$, as follows:

$$\Lambda_Q = \begin{pmatrix} \lambda_{11} & 0 & \cdots & 0 \\ \lambda_{21} & \lambda_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ \lambda_{Q1} & \lambda_{Q2} & \cdots & \lambda_{QQ} \end{pmatrix} \tag{9}$$

where the diagonal elements $(\lambda_{11}, \cdots, \lambda_{QQ})$ are all positive. We can obtain identified parameters to introduce this restriction. However, it is difficult to interpret the estimated factors directly from the obtained samples. Therefore, we rotate posterior samples and discuss them. The detailed procedures are available in section 3.4. The conditional posterior distributions are in the Appendix.

3.3. Number of Factors

One of the biggest problems in the factor analysis is how to determine the dimension of the factor. The most famous way is to determine from eigenvalues of the correlation matrix (Kaiser 1960; Guttman 1954). This method is called Kaiser Criterion, and suggests that the number of the eigenvalue over 1 is an appropriate dimension of the model. In addition, we can apply the likelihood of the model to selection. Especially in Bayesian models, we can use the marginal likelihood or DIC (Deviance Information Criterion). In addition, Lopes and West (2004) estimate the dimension of the model using the Reversible Jump algorithm developed by Green (1995).

In this study, we compare models using the Kaiser Criterion, marginal likelihood, and DIC. Note that eigenvalues are obtained from the correlation matrix of the observed variable.

3.4. Rotation

As mentioned above, we can rotate the factor loadings and scores. To rotate the factors, we can obtain useful managerial insights from the model. We also rotate the parameters using the following procedure.

At first, we calculate the rotation matrix R (where $R'R = I$). Let the h-th sample of the posterior distribution of the factor loadings parameter be $\Lambda^{(h)}$, where $h = 1, \cdots, H$. In addition, let the posterior mean be $\bar{\Lambda}$. Since we obtain the posterior mean value for each element, the samples and posterior mean have the following relationship: $\bar{\lambda}_{jq} = H^{-1} \sum_{h=1}^{H} \lambda_{jq}^{(h)}$. Next, we rotate the loadings for any criterion. In this study, we apply the varimax criterion. As a result, we can obtain the rotation matrix R.

The rotation affects the factor loadings, factor scores, and hierarchical parameters. Therefore, we also rotate these samples. Let $\lambda_j^{(h)}, f_i^{(h)}, \Delta^{(h)}$, and $\Gamma^{(h)}$ be the h-th posterior samples of each parameter. From these samples, the rotated samples are obtained by multiplying the rotation matrix as follows:

$$
\begin{aligned}
\lambda_j^{R(h)} &= R' \lambda_j^{(h)} \\
f_i^{R(h)} &= R' f_i^{(h)} \\
\Delta^{R(h)} &= R' \Delta^{(h)} \\
\Gamma^{R(h)} &= R' \Gamma^{(h)}
\end{aligned}
\tag{10}
$$

Using these rotated samples, we can calculate posterior mean, posterior standard deviation, HPD (highest posterior density) interval, and so on, to obtain useful insights.

4. Objective Products

In this study, we collect consideration set data from a consumer research survey. In this section, we propose the objective category, brands, and measures.

4.1. Overview of Objective Markets

In this study, we focus on the beer markets and 21 major brands, as shown in Table 1. The beer market is dominated by four manufacturers: Asahi, Kirin, Suntory, and Sapporo. The sum of the market share of these four firms accounts for over 98% of the market. All the targeted 21 brands are made by these firms. In addition, there are four submarkets in this market. In general, brands categorized as "premium beer" are relatively expensive, and "beer" follows next. Brands categorized as "low-malt" and "new genre" are relatively low-priced. Note that these four categories do not exactly correspond to the categorization in Japan's Liquor Tax Act. Strictly speaking, premium beers are categorized as beer. However, this categorization is commonly used by manufacturers and retailers. Although we can purchase 350ml cans, 500ml cans, and bottled beer in convenience stores, we focus on the 350ml purchase occasion. Therefore, we show 350ml can images in the questionnaire.

Table 1: Objective Brands and Submarkets

Manufacturer (Abbreviation)	Premium Beer (P)	Beer (B)	Low-Malt (L)	New Genre (N)
Kirin (K)		Lager Beer Ichiban Shibori	Tanrei Green Label	Sumikiri Nodogoshi Nama Koi Aji
Asahi (A)	Jukusen	Super Dry	Style Free	Asahi Off Clear Asahi
Suntory (S)	The Premium Malts	Malts		Kin Mugi Jokki Nama
Sapporo (P)	Yebisu Beer	Kuro Label	Hokkaido Namashibori	Kin No Off Hokkaido Premium Mugi To Hop

4.2. Consideration Set

As mentioned in Brisoux and Laroche (1980) and Shocker et al. (1991), the consideration set is the subset of the set of all available brands. The actual choice is based on this set. In the exclusive choice model, the number of chosen brands is only one. Suppose that the number of available alternatives is M and the size of consumer i's consideration set is C_i, the range of the set is $1 \leq C_i \leq M$. Note that there are some different kinds of choice subsets. Some previous studies propose a stepwise alternative selection model of consumer information processing. The consideration set is defined as one of the choice subsets within the process.

We ask respondents whether the brand is considered during the purchase. It is possible that the consumer chooses all available brands or none of the brands. We explicitly require respondents to choose "considered" brands. In the questionnaire, we show all 21 objective brands with brand names and package images. We ask respondents to check all brands that are considered during the purchase: *when you purchase a beer can, please check all brands that are considered during the purchase*. We collect binary answers; {0: do not consider, 1: consider}.

4.3. Consumer Attributes

As defined in the previous section, we try to explain factor scores by consumer characteristics such as demographic or psychographic attributes. Our model enables us to find relationships between each factor and consumer attributes.

In this study, we apply following five attributes: consumers' age, sex, objective knowledge (OK), subjective knowledge (SK), and familiarity (FM). The first two attributes are demographic aspects, while the latter three attributes are sub constructs of consumer knowledge. As Alba and Hutchinson (1987, 2000) pointed out, consumer knowledge of a product category consists of two sub constructs—familiarity and expertise. Further, the expertise is also divided into two parts—objective and subjective knowledge (Brucks 1985; Park and Lessig 1981).

Based on these studies, we define the effect of consumer knowledge as follows. The objective and subjective knowledge directly affect the factor score, while we define that familiarity moderates these knowledge constructs.

$$x_i = \{\text{Intercept}, \text{Age}, \text{Sex}, \text{OK}, \text{SK}, \text{OK} * \text{FM}, \text{SK} * \text{FM}\}' \tag{11}$$

where Age is the log of the observed age, and Sex is a binary value in which {0: male, 1: female}.

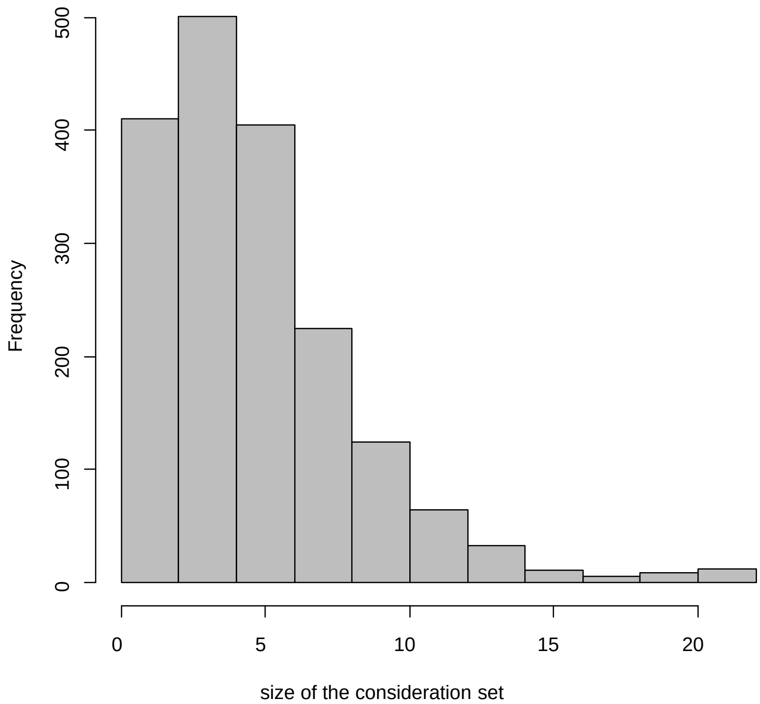
The measurement scales of the consumer knowledge constructs are listed in Appendix A.2. The familiarity, objective knowledge, and subjective knowledge scales are based on Moorman et al. (2004) and Brucks (1985).

5. Results

5.1. Data Collection

We collected data from October 11 to 15, 2013, through an internet survey. The 1,800 respondents in our sample were all over 20 years old. As a screening technique, we focus only on the consumers who drink beer. There are 1,093 males and 707 females. The average age is 46.8 years. Within the sample, the average size of the consideration set is 5.15, the mode is 4, the maximum value is 21 (all brands), and the minimum value is 1. Comparing with the previous study, Hauser and Wernerfelt (1990) report that the size of the consideration set in the beer category is between 3 and 7. Our result is also within this range. Figure 2 shows the histogram of the size of the consideration set.

Figure 1: Histogram of the Size of the Consideration Set



5.2. Model Comparison: Number of Submarkets

In this subsection, we examine the number of submarkets in the market. In other words, we determine the appropriate number of dimensions of the factor model. Table 2 shows the result of indicators obtained from each model. In the beer market, there are 5 eigenvalues that exceed 1. Therefore, from the Kaiser Criterion, the appropriate dimension is 5. In this study, we estimate 6 models, from the 1 factor model to the 6 factor model, and calculate the marginal likelihoods and DICs. The detailed procedures to obtain marginal likelihood and DIC are found in Newton and Raftery (1994) and Gelman et al. (2004), respectively. In addition, we also calculate the Bayes factors from the marginal likelihoods.

As Table 2 shows, although we examine models until 6 factors, the marginal likelihood and DIC of the 6 factor model is the highest. We additionally estimate the 7 factor model and find that the marginal likelihood is -10317.5 and the DIC is -13260.7. This suggests that the fitness of the model will improve as the number of factors increase. However, our purpose is to examine the submarkets. Although we need to interpret the results, it is difficult to interpret the result if the number of factors is too large. Therefore, in this study, we choose the 5 factor model based on the eigenvalue, and examine the results.

Table 2: Model Comparison Results

# of factors	Log Marginal Likelihood	Log Bayes Factor	DIC	Eigenvalue
1	-16090.0		-16849.2	4.01
2	-14083.2	2006.7	-15495.0	2.52
3	-12985.0	1098.2	-14858.3	1.52
4	-12077.5	907.5	-14341.5	1.49
5	-11212.0	865.5	-13789.7	1.10
6	-10648.7	563.4	-13287.7	0.95

Note: Log Bayes Factor of the k factor model is $m_k - m_{k-1}$ where, m_k is the log marginal likelihood of the k factor model.

5.3. Parameters

Table 3 shows the posterior mean of the factor loadings obtained from the 5 factor model. All the results are based on the rotated samples shown in the previous section. In Table 3, bold fonts indicate especially higher (lower) values among each factor. We find there are some submarkets in this market. For example, in the Factor 2 submarket, values of traditional and relatively high priced beer brands such as “Lager beer (K/B)” and “Ichian Shibori (K/B)” are negatively high. In contrast, the values of low calorie content and low price brands such as “Asahi Off (A/N)” and “Kin no Off (P/N)” are positively high. This suggests that in the Factor 2 submarket, consumers who have a positively high factor score tend to prefer relatively new low calorie content brands and avoid traditional brands. In contrast, the consideration probabilities of these traditional brands of consumers who have a negatively high score in this factor tend to be higher.

Similarly, values of low calorie content and low price brands such as “Asahi Off (A/N),” “Kin no Off (P/N),” and “Style Free (A/L)” are positively high in Factor 1. In Factor 3, the values of brands related to Hokkaido, “Hokkaido Premium,” and “Hokkaido Namashibori” are high. In

Factor 4, the premium beer brands called "The Premium Malts (S/P)" and "Yebisu Beer" have negatively higher scores. In Factor 5, brands that are classified as "nama (non pasteurized)" have higher scores, such as "Nodogoshi Nama (K/N)" and "Jokki Nama (S/N)." These brands also have the common property that they are both lower priced (new genre) brands.

From the factor loadings, we find that there are some submarkets. Based on this result, we further examine the relationships between the factor (submarket) preference and consumer attributes. In the next subsection, we examine the prior structure of the factor scores f_i.

Table 3: Factor Loadings

Brands (Manufacturer/Category)	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Tanrei Green Label (A/P)	1.02	−0.35	−0.04	0.22	0.19
Asahi Off (A/N)	**1.54**	**0.39**	−0.03	−0.10	0.32
Kuro Label (P/B)	−0.13	−0.73	0.71	−0.31	−0.10
Kin no Off (P/N)	**1.53**	**0.36**	0.65	0.03	0.54
The Premium Malts (S/P)	−0.21	−0.63	0.09	**−2.14**	0.13
Sumikiri (K/N)	0.36	−0.12	0.46	−0.10	0.65
Malts (S/B)	−0.03	−0.54	0.45	−0.76	0.09
Style Free (A/L)	**1.73**	0.11	0.11	0.18	0.30
Hokkaido Premium (P/N)	0.30	−0.06	**1.93**	−0.25	0.58
Lager Beer (K/B)	−0.11	**−1.65**	0.11	−0.08	0.07
Kin Mugi (S/N)	0.16	−0.01	0.21	−0.22	0.95
Ichiban Shibori (K/B)	0.10	**−1.32**	0.14	−0.39	0.05
Nodogoshi Nama (K/N)	0.33	−0.47	−0.08	**0.36**	**1.27**
Hokkaido Namashibori (P/L)	0.33	−0.34	**1.83**	−0.03	0.54
Koi Aji (K/N)	0.88	0.05	0.48	−0.09	0.36
Yebisu Beer (P/P)	−0.24	−0.69	0.49	**−0.88**	**−0.17**
Mugi to Hop (P/N)	0.11	0.02	0.69	−0.17	0.66
Super Dry (A/B)	−0.06	−0.42	−0.01	−0.11	0.11
Jokki Nama (S/N)	0.17	−0.11	0.28	0.04	**1.27**
Jukusen (A/P)	0.26	−0.12	0.87	−0.63	0.05
Clear Asahi (A/N)	0.36	0.06	0.12	−0.01	0.76

Note: For manufacturers, A: Asahi, K: Kirin, S: Suntory, and P: Sapporo. For categories: P: Premium Beer, B: Beer, L: Low-malt beer, and N: New Genre. Bold fonts indicate especially higher (lower) values among each factor.

5.4. Consumer Characteristics and Factors

Our model assumes the prior structure on each consumer's factor scores f_i. In this study, we assume that consumer characteristics are explained by their demographic traits (age and sex) and knowledge (objective knowledge, subjective knowledge, and familiarity). Table 4 shows the result of the parameter Δ. In this table, underlined fonts indicate that 0 lies outside the 10% HPD (highest probability density) intervals of the parameter. Further, bold fonts indicate that 0 lies outside the 5% HPD intervals. The detailed procedure to obtain HPD intervals is found, for example, in Chen et al. (2000).

Table 4 Result of Prior Parameter $\boldsymbol{\Delta}$

Δ	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Intercept	0.79	$\mathbf{1.26}$	-0.07	0.28	$\mathbf{1.28}$
Sex	$\mathbf{0.26}$	$\mathbf{0.18}$	0.08	-0.09	0.10
Age	$\mathbf{-0.32}$	$\mathbf{-0.30}$	-0.16	0.04	$\mathbf{-0.32}$
SK	$\underline{0.61}$	0.30	$\mathbf{1.07}$	-0.30	-0.08
OK	0.20	$\mathbf{-0.68}$	$\mathbf{1.02}$	$\mathbf{-1.02}$	$\underline{-0.60}$
SK*FM	0.02	-0.59	-0.74	-0.64	-0.27
OK*FM	0.06	0.45	0.08	0.60	$\mathbf{1.60}$

Note: SK: Subjective Knowledge, OK: Objective Knowledge, FM: Familiarity. Underlined fonts indicate that 0 lies outside the 10% HPD interval, and Bold fonts indicate that 0 lies outside the 5% HPD interval.

To assume the consumer attribution for the prior structure of the factor scores f_i, we can find useful implications of consumer behaviors for each submarket. For example, we find that in Factor 2, if f_{i2} is positively higher, the consumer tends to consider low calorie contained brands, while if f_{i2} is negatively higher, the consumer prefers traditional well-tasted brands in Table 3. Based on this, we can interpret the Table 4 that the result implies that younger female consumers tend to consider low calorie content brands and older male customers consider traditional brands. In addition, we find that consumers who have higher objective knowledge tend to prefer traditional brands.

Similarly, in Factor 1, we find that younger female consumers tend to consider low calorie content new genre brands. In addition, we find from Factor 3 that consumers who have higher objective knowledge and subjective knowledge tend to consider low-malt or new genre "Hokkaido" related brands. In Factor 4, premium beer category brands are considered by consumers who have higher objective knowledge. Finally, in Factor 5, "Nama (non pasteurized)" brands are considered by consumers who are young and have high familiarity and subjective knowledge. Since the construct of familiarity is measured by the frequency of use and purchase, and subjective knowledge is measured by the confidence of consumers' own knowledge, this result implies that consumers who drink beer and have more confidence in their knowledge than ordinary people tend to prefer these "Nama" brands.

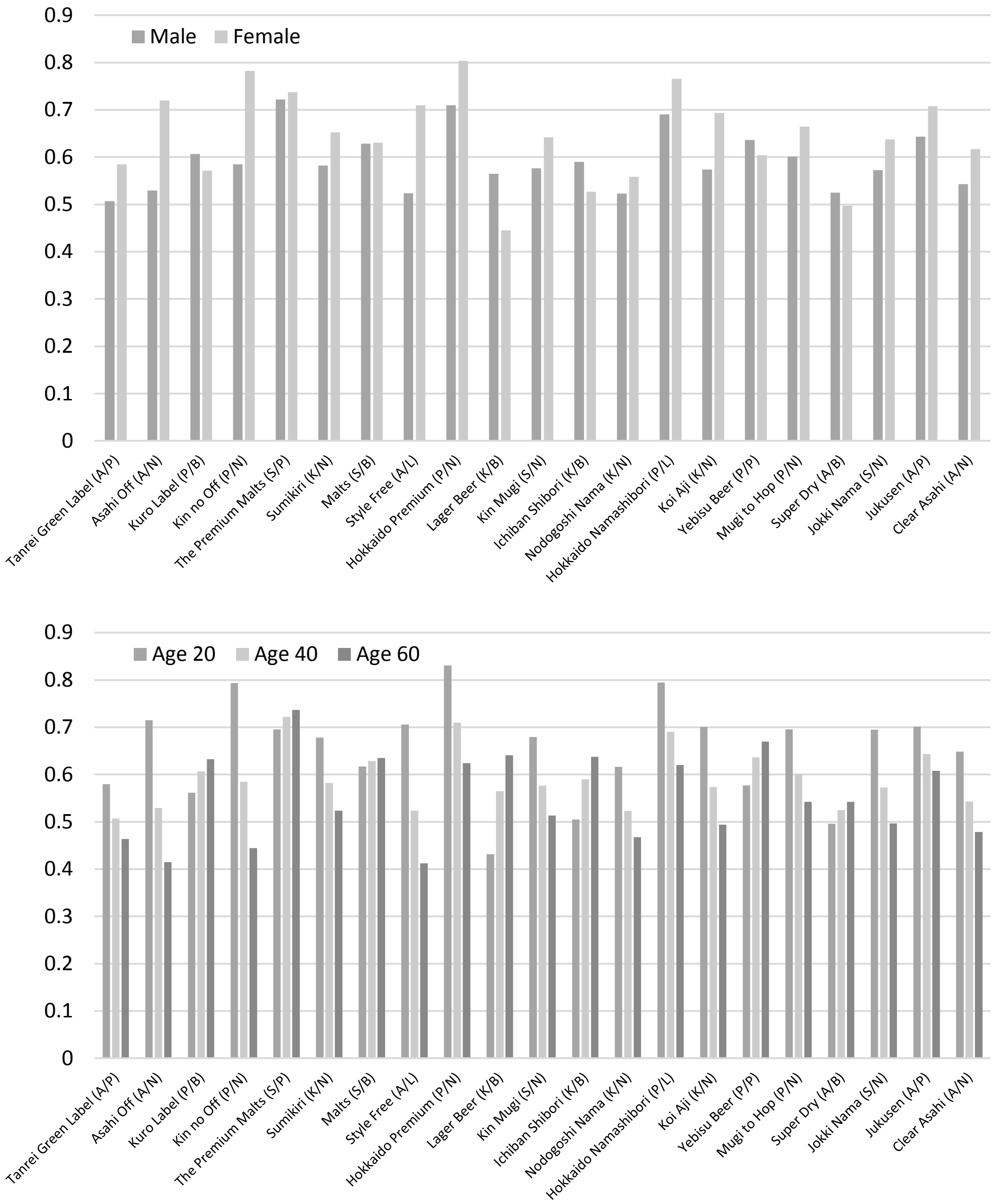
As a summary, our model enables managers not only to identify the submarket structure but also to show the relationship between each submarket and consumer attribution such as demographic aspects and knowledge.

5.5. Consumer Attribution and Consideration Probability

Since our model is based on the probabilistic choice model, we can estimate the consideration probability using the factor loadings Λ and prior variable of the factor scores. Suppose that a consumer whose attribution is $\tilde{w}$, the consideration probability of the consumer is obtained from $\Phi(\Lambda\Delta\tilde{w})$, where $\Phi(x)$ is the standard normal distribution function evaluated on x.

The estimated considered probability is shown in Figure 3. The upper figure shows the difference between male and female consumers. In this figure, other attributes are as follows: 40 years old, SK=OK=FM=0.5. This suggests that female consumers tend to prefer low calorie brands and male customers prefer beer category (B) brands. The lower figure of Figure 3 shows that younger customers tend to consider new genre (N) and low-malt (L) category brands, while older customers prefer beer category (B) brands.

Figure 2: Difference in the Consideration Probability (Upper: Sex, Lower: Age)



Note: The upper figure shows the probabilities of 40 years old consumers. The lower figure shows the probabilities of male consumers.

6. Conclusion

In this study, we focus on the competitive market, which has many brands, and identify the submarket structure using the discrete choice factor analysis model. This study makes two contributions, as follows.

First, our study identifies the consideration set based submarket structure. Since we can scarcely observe the consideration set of consumers, the number of studies related to the consideration set formulation and its component are relatively few compared with actual choice (purchase) studies. Our findings shed light on the field.

Second, we apply the hierarchical factor analysis model to examine the market structure. Our model allows the binomial objective variable and enables us to estimate probabilities. This is feasible to examine the choice problem and we succeed in applying the model to the consideration probability. Further, we assume the prior linear combination mean on the factor scores f_i. This enables us to identify the relationships between submarkets and consumer characteristics. Further, to rotate the posterior samples, we can easily interpret the result and obtain useful implications about the market.

As issues for future research, we propose two points. First, we have to incorporate the model as part of the conditional choice models proposed in Bucklin et al. (1998) and van Heerde and Neslin (2008). In this study, we only focus on the consideration set composition. However, our goal is to estimate the choice probability of each brand. Second, we have to find a criterion to decide the appropriate number of factors. In this study, the marginal likelihood and DIC did not work well. Although these indicators tend to prefer large dimensional models, these are difficult to interpret for managers and analysts. We have to find a criterion to solve the trade-off.

References

Alba, J. W. and J. W. Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13, 411-454.

Alba, J. W. and J. W. Hutchinson (2000), "Knowledge Calibration: What Consumers Know and What They Think They Know," *Journal of Consumer Research*, 27, 123-156.

Albert, J. H. and S. Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669-679.

Andrews, R. L. and T. C. Srinivasan (1995), "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32, 30-41.

Ansari, A. and K. Jedidi (2000), "Bayesian Factor Analysis for Multilevel Binary Observations," *Psychometrika*, 65, 475-496.

Brisoux, J. E. and M. Laroche (1980), "A Proposed Consumer Strategy of Simplification for Categorizing Brands," in *Evolving Marketing Thought for 1980*, ed J. D. Summery and R. D. Taylor, Southern Marketing Association, 112-114.

Brucks, M. (1985), "The Effects of Product Class Knowledge on Information Search Behavior," *Journal of Consumer Research*, 12, 1-16.

Bucklin, R. E., S. Gupta, and S. Siddarth (1998), "Determining Segmentation in Sales Response across Consumer Purchase Behaviors," *Journal of Marketing Research*, 35, 189-197.

Chen, M. H., Q. M. Shao, and J. G. Ibrahim (2000), *Monte Carlo Methods in Bayesian Computation*, New York: Springer.

Chiang, J., S. Chib, and C. Narasimhan (1999), "Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity," *Journal of Econometrics*, 89, 223-248.

Elrod, T. and M. P. Keane (1995), "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data," *Journal of Marketing Research*, 32, 1-16.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004), *Bayesian Data Analysis, Second Edition*, Chapman and Hall/CRC.

Geweke, J. F. and G. Zhou (1996), "Measuring the Pricing Error of the Arbitrage Pricing Theory," *The Review of Financial Studies*, 9, 557-587.

Green, P. J. (1995), "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," *Biometrika*, 82, 711-732.

Guttman, L. (1954), "Some necessary conditions for common factor analysis," *Psychometrika*, 19, 194-162.

Hauser, J. R. and B. Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 17, 74-81.

Howard, J. A. and J. N. Sheth (1969), *The Theory of Buyer Behavior*, John Wiley and Sons, Inc.

Kaiser, H. F. (1960). "The Application of Electronic Computers to Factor Analysis," *Educational and Psychological Measurement*, 20, 141-151.

Lopes, H. F. and M. West (2004), "Bayesian Model Assessment in Factor Analysis," *Statistica Sinica*, 14, 41-67.

Manski, C. (1977), "The Structure of Random Utility Models," *Theory and Decision*, 8, 229-254.

Moorman, C., K. Diehl, D. Brinberg, and B. Kidwell (2004), "Subjective Knowledge, Search Locations, and Consumer Choice," *Journal of Consumer Research*, 31, 673-680.

Newton, M. A., and A. E. Raftery (1994). Approximate Bayesian Inference with the Weighted Likelihood Bootstrap, *Journal of the Royal Statistical Society. Series B*, 64, 3-48.

Nierop, E. V., B. Bronnenberg, P. Paap, M. Wedel, and P. H. Franses (2010), "Retrieving Unobserved Consideration Sets from Household Panel Data," *Journal of Marketing Research*, 47, 63-74

Park, C. W. and V. P. Lessig (1981), "Familiarity and Its Impact on Consumer Decision Biases and Heuristics," *Journal of Consumer Research*, 8, 223-230.

Roberts, J. H. and J. M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28, 429-440.

Roberts, J. H. and J. M. Lattin (1997), "Consideration: Review of Research and Prospects for Future Insight," *Journal of Marketing Research*, 34, 406-410.

Shocker, A. D., M. Ben-Akiva, B. Boccara, and P. Nedungadi (1991), "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions," *Marketing Letters*, 2, 181-197.

Siddarth, S., R. E. Bucklin, and D. G. Morrison (1995), "Making the Cut: Modeling and Analyzing Choice Set Restriction in Scanner Panel Data," *Journal of Marketing Research*, 32, 255-266.

Simon, H. A. (1947), *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, New York: Macmillan.

van Heerde, H. J. and S. Neslin (2008), "Sales Promotion Models," in *Handbook of Marketing Decision Models*, ed. B. Wierenga, Springer, 107-162.

A. Appendix

A.1. Posterior Distributions

For the latent variable z_{ij}^*, we can generate samples from the following truncated normal distribution:

$$z_{ij}| \cdot \sim \begin{cases} TN_{(0,\infty)}(\alpha_j + \lambda_j' f_i \,, 1) \\ TN_{(-\infty,0)}(\alpha_j + \lambda_j' f_i \,, 1) \end{cases} \quad \text{(A.1)}$$

The sample of the intercept parameter α_j is obtained from the following univariate normal distribution:

$$\alpha_j | \cdot \sim N(m_j, s_1^2) \quad \text{(A.2)}$$

where $s_1^2 = (s_0^{-2} + N)^{-1}$ and $m_j = v_j^2 \left(\sum_{i=1}^N \left(z_{ij}^* - \lambda_j' f_i \right) + s_0^{-2} a_0 \right)^{-1}$.

The factor loadings Λ are obtained for each row λ_j. However, we have to consider the first Q rows. Therefore, in this section, we will explain the first to Q-th row and Q-th to J-th row separately.

At first, in the first to Q-th row, λ_j is constrained ($1 \le j \le Q$); that from the $(j+1)$-th element to the Q-th element is 0. Additionally, the j-th element is constrained to be positive. Considering these conditions, the posterior distribution is following a constrained multivariate normal regression.

$$\lambda_{j,1:j} | \cdot \sim N_j(\mu, \Psi) \mathbf{1}(\lambda_{jj} > 0) \quad \text{(A.3)}$$

where $\Psi = \left(V_{0j}^{-1} + E_j' E_j \right)^{-1}$, $\mu = \Psi \left(V_{0j}^{-1} g_{0j} + E_j' Y_j \right)^{-1}$, $V_{0j} = v_0^2 I_j$, $g_{0j} = g_0 \mathbf{1}_j$ ($\mathbf{1}_j$ is a j-dimensional 1 vector), $E_j = F_{\cdot,1:j}$, and $Y_j = Z_{\cdot,1:j}^*$. We divide this posterior distribution into two parts, from the first to $(j-1)$-th element $\lambda_{j,1:(j-1)}$ and j-th element λ_{jj}. Hereafter, let $j^* = 1:(j-1)$.

$$\begin{aligned} \lambda_{j,j^*} | \lambda_{jj}, \cdot &\sim N_j(\mu^*, \Psi^*) \\ \lambda_{jj} | \lambda_{j,j^*}, \cdot &\sim TN_{(0,\infty)}\left(\mu^\dagger, \Psi^\dagger\right) \end{aligned} \quad \text{(A.4)}$$

where $\mu^* = \mu_{j^*} + \Psi_{j^*,j} \Psi_{jj}^{-1} \left(\lambda_{jj} - \mu_j \right), \Psi^* = \Psi_{j^*,j^*} - \Psi_{j^*,j} \Psi_{jj}^{-1} \Psi_{j,j^*}$, and $\mu^\dagger = \mu_j + \Psi_{j,j^*} \Psi_{j^*,j^*}^{-1} \left(\lambda_{j^*} - \mu_{j^*} \right), \Psi^\dagger = \Psi_{jj} - \Psi_{j,j^*} \Psi_{j^*,j^*}^{-1} \Psi_{j^*,j}$.

While for $(Q+1)$-th to J-th row of the factor loading, there are no constraints. Therefore, we can generate samples from the following multivariate normal distribution.

$$\lambda_j | \cdot \sim N_J(\mu, \Psi) \tag{A.5}$$

where $\Psi = \left(V_{0J} + F'F\right)^{-1}$, $\mu = \Psi\left(V_{0J} g_{0J} + F'Z^*\right)^{-1}$, $V_{0J} = v_0^2 I_J$, and $g_{0j} = g_0 \mathbf{1}_J$.

Posterior samples of factor scores are obtained from the following Q-dimensional multivariate normal distribution.

$$f_i | \cdot \sim N_Q\left(\left(I_Q + \Lambda'\Lambda\right)^{-1}(\Delta w_i + \Lambda' z_i^*), \left(I_Q + \Lambda'\Lambda\right)^{-1}\right) \tag{A.6}$$

The sample of parameter δ_p, which affects the p-th column of the factor score $f_p = \left(f_{1p}, \cdots, f_{Np}\right)'$, is generated from the following K-dimensional multivariate normal distribution.

$$\delta_p | \cdot \sim N_K\left((\Sigma_0^{-1} + W'W)^{-1}\left(\Sigma_0^{-1} d_0 + W' f_p\right)^{-1}, (\Sigma_0^{-1} + W'W)^{-1}\right) \tag{A.7}$$

where $W = (w_1, \cdots, w_N)'$.

A.2. Measures

Table A.1 Measurement Scales

Familiarity (2 items, 7-point scale)
Reliability: 0.893
Directions: I would like to ask you about canned beers (or alcohol that tastes like beer). Please answer the following questions.

Items	Anchors
1) How often do you purchase canned beers? 2) How often do you drink canned beers?	1) less than one can a month 2) one can a month 3) few (2 to 3) cans a month 4) one can a week 5) few (2 to 3) cans a week 6) one can a day 7) more than two cans a day

Subjective Knowledge (3 items, 7-point scale)
Reliability: 0.958
Directions: Please answer the following questions.

Items	Anchors
1) Rate your knowledge of beer product information compared to the average consumer. 2) Rate your confidence in using beer product information compared to the average consumer. 3) I feel confident about my ability to comprehend beer information on product labels.	1) much less 2) less 3) a little less 4) average 5) a little more 6) more 7) much more

Objective Knowledge (21 items, binary scale)
Reliability: 0.804
Directions: For the following canned beer brands, match the category and manufacturer to the brand.

Items	Anchors
1) Jukusen (1) 2) The Premium Malts (5) 3) Yebisu (5) 4) Super Dry (1) 5) Ichiban Shibori (3) 6) Ragger Beer (3) 7) Malts (5) 8) Kuro Label (7) 9) Style Free (2) 10) Tanrei Green Label (4) 11) Hokkaido Namashibori (8)	1) beer brand manufactured by Asahi 2) low malt or new genre brand manufactured by Asahi 3) beer brand manufactured by Kirin 4) low malt or new genre brand manufactured by Kirin 5) beer brand manufactured by Suntory 6) low malt or new genre brand manufactured by Suntory 7) beer brand manufactured by Sapporo 8) low malt or new genre brand manufactured by Sapporo

12) Asahi Off (2) 13) Clear Asahi (2) 14) Koiaji (4) 15) Sumikiri (4) 16) Nodogoshi Nama (4) 17) Kin Mugi (6) 18) Jokki Nama (6) 19) Kin no Off (8) 20) Mugi to Hop (8) 21) Hookkaido Premium (8) (Correct answers are in brackets)	