数理解析研究所講究録
第 1954 巻 2015 年 104-124

Are manufacturers' efforts to improve their brands' reputation really rewarded? The case of Japanese yogurt market

Tomohito KAMAI and Yuichiro KANAZAWA
Division of Policy and Planning Sciences, University of Tsukuba

1 Introduction

Since yogurt was first introduced in Japan in 1950's, the market became one of the main category for the grocery retail channel with the second largest sales in food category according to the national retail survey conducted between September 2012 to February 2013.[1]. The recent market growth is said to be stimulated by a group of products with newly found lactic-acid bacilli, which are claimed to enhance immune strength and prevent consumers from virus-infection, allergies and so forth. The traditional marketing theory would predict that these manufacturers' efforts are rewarded with high margins. The average price of yogurt, however, kept decreasing over the last decade and the temporal price reduction ("TPR" henceforth) is prevalent practice in this category, with 66.7% of supermarket engaged in TPR in a sampled week according to the retail survey in 2007.[2]

In order to answer if the manufacturers are really rewarded for their innovations, we employ a framework of [4] whereby retail prices are decomposed into manufacturers' and retailers' margins and marginal cost to assess the relative magnitude of them. In addition to strategic interaction among manufacturers and retailers and consumer state dependence, a model is able to accommodate forward-looking pricing policy of firms as description of firms' behavior and margins each firm obtains as a result of their behavior could be drastically altered if we fail to model such behavior. In the research of [4], for example, both manufacturers and retailers in U.S. cereal market are shown to set prices accounting for the effect of current prices on future profit. In this research, we apply the model to the yogurt data in Japanese market to correctly answer our inquiry with the most plausible framework to describe the market.

[1] *Source:KSP-POS Market Trend Report, vol.48*, Knowledge on Sales Promotion Service Providers.

[2] Compared to 2003, the average retail price of boxed yogurt in stores in Tokyo area fell by 7.5% in 2008, and it further fell by 14.9% in 2013 according to "Retail Survey" conducted by Statistics Bureau, Ministry of Internal Affairs and Communications, Japan. Data regarding TPR are obtained from "National Survey of Prices" conducted by Statistics Bureau, Ministry of Internal Affairs and Communications and calculated from data of "Distribution of Regular Prices and Sale Prices by Sales Floor Space, Type of Outlets - Japan, City Groups, Prefectures."

The rest of paper is organized as follows. The next section describes the model. In section 3, we present our estimation procedure. We briefly explain our data in section 4. In section 5, we will present and discuss results for empirical analysis. Section 6 concludes.

2 The Model

In this section, we specify both demand- and supply-side models. As we implied, there are three major dimensions in the modeling framework, which are strategic interaction among manufacturers and retailers, consumer state dependence, and forward-looking behavior of firms. Out of them, consumer state dependence is specified in demand-side behavior and the rest is specified in supply-side behavior. This approach of structural market equilibrium model enables the analysis of supply-side behavior by observing only the demand-side data, which is an advantage of the model as supply-side information is rarely available to researchers.

The model is widely used in the literature as it offers rich insights to marketing issues. The main use of the model includes theory testing, what-if analysis, and identification of the determinant of marketing power and profitability among channel members [10]. As for theory testing, for example, [7] justify a policy of uniform pricing where different items under the same brand name have identical prices in spite of the difference in some product attribute such as flavors in yogurt category. By identifying the competitive structure of the market and the source of the profitability of participants in the same distribution channels, each participant can figure out how they can efficiently align their marketing mix options to achieve maximum return given their competitive environment. In this line, [2] use the model to calibrate the monetary value of target pricing and [9] investigate the impact of brand positioning and change in price for cars under Bertrand competition. [8] investigate the relationship between the retail environment and intensity of manufacturer condition. [16] investigate the effect of new brand introduction to competitive relationships between firms. Recently, the power balance between manufacturers and retailer is often discussed when retailers are armed with their store brands which have multiple effects such as increased bargaining power with respect to manufacturers, inducing store traffic and building store loyalty in the context of among retailers competition and so forth. The examples in this line include [5] and[13]. The other examples of papers in this line include [3], [17], [20], [19] [7] and [4] to name a few.

2.1 Demand–Side Specification

Because supply-side behavior is estimated conditional on the estimation results of demand-side model, we start with demand-side model.

2.1.1 The brand choice model

Let us suppose there are $j = 1, \dots, J$ brands in the market and each household $i = 1, \dots, I$ has $t_i = 1, \dots, T_i$ purchasing occasions. We employ the multinomial logit model for household brand choice behavior with the latent class model to accommodate the heterogeneity across households [11]. Specifically, the deterministic part of the utility of household i choosing brand j at its t_i-th purchasing occasion is defined as

$$v_{ijt_i} = \boldsymbol{x}_{jt_i} \cdot \boldsymbol{\beta}_s + \mathrm{sim}_{kj} \cdot SD_s + \xi_{jt_i} \tag{1}$$

where the outside option is expressed as $v_{i0t_i} = 0$ and where vector $\boldsymbol{x}_{jt_i}$[3] includes brand dummy variables and price of brand j a household i faces on purchasing occasion t_i, sim_{kj} is the attribute similarity index for brand j with respect to the previously purchased brand k, and ξ_{jt_i} is the unobserved demand characteristics which can be observed by firms and households but not by a researcher. The examples of unobserved demand characteristics are national advertisement, coupon availability, shelf space allocations and so forth. As prevalent in this study field, we assume it commonly affects all households [3, 18, 19]. It is empirically well known that ignoring unobserved product characteristics leads to a biased estimate of price effect as they could be correlated with prices [1, 18, 3, 2, 14, 19]. To avoid this problem, we employ an idea of two-stage least squares. Parameters to be estimated are $\boldsymbol{\beta}_s$ and SD_s, where a subscript $s = 1, \dots, S$ corresponds to segment (i.e., a subset to which households belong to, where those in the same segment are assumed to be the same in terms of responsiveness to marketing mix variables).

2.1.2 The attribute similarity index

We use the attribute similarity index to express the state dependence in household brand choice behavior, following [4].[4] In their specification, each brand is allocated with a set of attributes by a researcher. Each attribute has different levels, and brands are

[3] The term ξ_{jt_i} is a subset of ξ_{jt} where the latter is defined for all calendar dates and brands in the panel, and the former is retrieved from the latter according to t_i. On the other hand, the values of $\boldsymbol{x}_{jt_i}$ may be different depending on households even when two households shop at the same time as temporal price reduction such as coupon may only be available to a specific household.

[4] The idea of the attribute similarity index can be found in previous papers (e.g., [12]), but the specification in previous literature requires questionnaire which explicitly asks subjects for the perceived similarity between listed brands. The advantage of the specification of [4] is that it does not require such information and similarity between brands can be calibrated from the data, although the level of attributes shared by brands must be set by researchers.

assumed to be similar if they share the same level of attributes. The degree of similarity between brands increases with the number of attribute levels shared by these brands.

Employing the attribute similarity index enables a researcher to examine how each brand attribute contributes to the perception of similarity between brands among consumers. Apparently, this approach would yield richer insight on consumer brand choice behavior and on brand positioning compared to the prevalent approach such as employing the lagged brand indicator variable. Specifically, the similarity between the brand purchased on the previous occasion (brand k) and the brand a household faces on the current purchase occasion (brand j) is specified as

$$\text{sim}_{kj} = \frac{I_{kj} + \sum_{p=1}^{P} I_{kjp} \cdot r_p}{1 + \sum_{p=1}^{P} r_p}, \tag{2}$$

where I_{kj} is an indicator variable taking unity if $k = j$, I_{kjp} is an indicator variable taking unity if two brands share the same level of attribute $p = 1, \cdots, P$, and $r_p > 0$ is importance weight associated with attribute p to be estimated. As (2) implies, the similarity index is designed to take value between 0 (brands are totally dissimilar) and 1 (brands are identical). The parameter of the attribute similarity index, SD_s, can either be positive or negative which corresponds to inertia (i.e., a previous brand consumption experience raises the probability of repurchasing a brand) and variety-seeking (i.e., a previous brand consumption experience lowers the probability of repurchasing a brand) respectively. Following [4], we specify SD_s to be the function of demographic variables as

$$SD_s = \gamma_{s0} + \mathbf{D}_i \cdot \gamma_s \tag{3}$$

where $\mathbf{D}_i$ is vector of demographic characteristics of household i, γ_{s0} is an intercept term, and γ_s is vector of parameters for $\mathbf{D}_i$. The available demographic information in our panel is gender and age.

2.2 Supply–Side Specification

Following the preceding research, we assume that the retailer is a local monopolist which maximizes its joint category profit.[5] The assumption of a local monopolist is often justified by empirical reports which find that there is little evidence of among store competitions [3, 17, 19, 4]. We further assume that there are multiple manufacturers which sell their brands through a common retailer. Manufacturers are allowed to produce multiple brands.

After estimating demand side parameters, we will estimate the margins of manufacturers and a retailer under four different games, which arise from the combination of two games in

[5] A retailer could use the other pricing rules such as brand profit maximization where it sets up a profit function for each brand. However, [17] empirically shows that a retailer attains a maximum profit when it engages in category profit maximization, which supports the assumption widely adopted in the literature.

horizontal strategic interaction among manufacturers and two games in vertical strategic interaction between manufacturers and a retailer. Two games in horizontal strategic interaction are Bertrand competition and tacit collusion, where Bertrand competition refers to own-brands profit maximizing behavior of each manufacturer and tacit collusion refers to the behavior of manufacturers which collectively maximize total profit from all brands in the market. Two games in vertical strategic interaction are manufacturer Stackelberg and vertical Nash. In the manufacturer Stackelberg game, manufacturers act as Stackelberg leaders with respect to a retailer and choose their wholesale prices anticipating a reaction from a retailer and wholesale prices of competing brands. In this case, the retailer chooses retail prices to maximize its profit taking wholesale prices as given. In the vertical Nash game, manufacturers and a retailer move simultaneously; they choose prices anticipating the profit maximizing behavior of the others [6, 17]. We reserve the derivation of margins in Appendix. Our derivation much follows [19] and [4].

After calculating margins of manufacturers and a retailer, we will estimate marginal cost of each brand using variables such as prices of ingredients. Finally, we will calculate likelihood for each model and game, and compare the results by Vuong test statistics.

3 Estimation

3.1 Demand–Side Estimation

3.1.1 Pricing equation

As prices may be correlated with unobserved demand characteristics, we first set up the pricing equation

$$p_{jt} = \kappa_0 + z_{jt} \cdot \kappa_1 + \eta_{jt} \tag{4}$$

where z_{jt} is an instrument which is correlated with p_{jt} but not with ξ_{jt}, κ_0 and κ_1 are parameters to be estimated, and η_{jt} is a random error term. Note that this equation is defined for calendar date $t = 1, \ldots, T$. We estimate $\widehat{p_{jt}}$ and $\widehat{\eta_{jt}}$ by ordinary least squares.

Next, ξ_{jt} is obtained as residual in the following equation:

$$\ln \tilde{S}_{jt} - \ln \tilde{S}_{0t} = \boldsymbol{x}_{jt} \cdot \boldsymbol{\beta} + \mathrm{sim}_{kj} \cdot SD + \xi_{jt} \tag{5}$$

where $\ln \tilde{S}_{jt}$ and $\ln \tilde{S}_{0t}$ are the log of observed market shares of brand j and outside good at time t respectively.

If price endogeneity exists, the terms ξ_{jt} and η_{jt} will be correlated.[6] This correlation should arise as η_{jt} can represent both demand and cost shock (i.e., if the unobserved

[6]As $\kappa_0 + z_{jt} \cdot \kappa_1$ is uncorrelated with ξ_{jt} by construction, η_{jt} represents a correlated (with ξ_{jt}) part of p_{jt}.

demand characteristic is desirable, it is reasonable to assume it incurs cost). In order to check the existence of price endogeneity, we assume that ξ_{jt} and η_{jt} jointly follow the bivariate normal distribution as correlation in that distribution equates dependence between them. We also assume that their means are both zero, and their moments exist up to the second order.

3.1.2 Likelihood function

The likelihood of purchase history of household i is written as

$$L_i = \prod_{t_i=1}^{T_i} \int \left\{ \prod_{j=0}^{J} [\mathrm{Pr}_{ijt_i}]^{y_{ijt_i}} \times f(\xi_{jt_i}|\eta_{jt_i}) \times f(\eta_{jt_i}) \right\} d\xi_{jt_i} \tag{6}$$

where y_{ijt_i} is an indicator function taking unity if household i chooses brand j at time t and 0 otherwise, $f(\xi_{jt}|\eta_{jt})$ is the conditional density of ξ_{jt}, and $f(\eta_{jt})$ is the density function of η_{jt}. Similarly to ξ_{jt_i}, the term η_{jt_i} is a subset of η_{jt}, which is defined for all calendar dates in the panel. In this paper, we employ the latent class model under which the likelihood function as in (6) for household i is replaced with $L_i(S_i = s)$, the likelihood of household i belonging to the segment s or $S_i = s$. Then we have the likelihood for whole panel data as

$$L = \prod_{i=1}^{I} \left\{ \prod_{s=1}^{S} L_i(S_i = s) \times \mathrm{Pr}_i(s) \right\} \tag{7}$$

where S is the number of segments and $\mathrm{Pr}_i(s)$ is the membership probability to segment s of household i. Parameters $\boldsymbol{\beta}_s$ and SD_s are estimated by maximizing this likelihood function.

3.2 Supply–Side Estimation

3.2.1 Marginal cost

We specify the marginal cost equation as

$$mc_{jt} = w_{j0} + \mathbf{input}_{jt} \cdot \boldsymbol{w}_r \tag{8}$$

where w_{j0} is a brand-specific intercept term, $\mathbf{input}_{jt}$ is vector of observable cost shifters, and $\boldsymbol{w}_r$ is corresponding vector of parameters. For the notational convenience, let $\boldsymbol{w} \equiv (w_{j0}, \boldsymbol{w}_r)$. Now to estimate $\boldsymbol{w}$, we utilize the following equation

$$p_{jt} - \widehat{CMM}_{jt} - \widehat{CMR}_{jt} \;=\; mc_{jt} + \varepsilon_{jt} \tag{9}$$

where $\widehat{CMM}_{jt}$ and $\widehat{CMR}_{jt}$ are computed margin of manufacturers and a retailer for brand j at time t respectively, and ε_{jt} is a random error term. Assuming the error term ε_{jt} follows a normal distribution with mean zero and finite variance (which is to be estimated), the right-hand side of the equation

$$\varepsilon_{jt} \quad = \quad p_{jt} - \widehat{CMM}_{jt} - \widehat{CMR}_{jt} - w_{j0} - \mathbf{input}_{jt} \cdot \boldsymbol{w}_r \tag{10}$$

also follows the normal distribution. Then we have the likelihood function of the supply-side as

$$\prod_{t=1}^{T}\prod_{j=1}^{J} g(\varepsilon_{jt}) \tag{11}$$

where $g(\cdot)$ is the marginal density of ε_{jt}, to estimate $\boldsymbol{w}$ and to calculate Vuong test statistics.

4 Data

We use scanner-panel data of yogurt purchases from anonymous retail chain in the western Tokyo in January 2007 to December 2008. Between two type yogurts – box type and snack type – we chose the latter type for our empirical analysis as the former type may also be used for cooking. Out of brands remained on sale throughout the period, we chose 7 brands which had enough purchasing records across stores, as we would like to use the average yogurt prices in these stores as instruments for prices of yogurt in particular store we would analyze.[7] After choosing households who only purchased the selected 7 brands at least twice, 183 households who made 15,194 shopping trips and 2,550 yogurt purchases remained. In the data, 76.5% of purchases were made by a female member of household. The average age of consumers in the panel is 59.4 with standard deviations of 19.6. The minimum and maximum ages of consumers in the panel are 14 and 94 respectively.

The summary of brands is summarized in Table 1. The attributes we used for the attribute similarity index were "Raw milk usage" (the proportion of raw milk in yogurt, 3 levels), "Fat level" (the fat amount contained, 3 levels) and "Ager usage" (whether yogurt contains ager or not, 2 levels). Ager is used to produce so called "hard-type" yogurt, which has texture like pudding unlike plain-type yogurt. Out of these brands, we are especially interested in brand 3, 5, and 6; brand 3 is differentiated in terms of taste (it is the only brand using only raw milk), brand 5 is the yogurt with special lactic-acid

[7] The other stores had at least 20 dates without a single sale of any brands during two years. We chose to exclude them from our analysis, as brand switch could have been attributed to the fact that some of them were out-of-stock in these stores. In this paper, we are not focusing on this kind of forced brand switching behavior.

Table 1: Summary of brands.

	Average price (yen per gram)	Manufacturer ID	Market share	Raw milk usage	Fat level	Ager usage	Fat content (g/100g)	Sugar content (g/100g)
Brand 1	0.459	1	1.14%	No	Middle	Yes	2.47	7.77
Brand 2	0.486	2	2.95%	Partial	Middle	Yes	2.05	14.6
Brand 3	0.488	3	0.86%	All	High	Yes	4.10	14.9
Brand 4	0.483	4	1.08%	Partial	Low	No	1.76	15.2
Brand 5	1.113	4	3.22%	Partial	Middle	No	3.04	9.73
Brand 6	1.113	4	1.35%	Partial	Low	No	1.43	9.20
Brand 7	0.834	5	2.31%	Partial	Low	No	1.88	13.4

bacilli, and brand 6 is a low fat version of brand 5. To compare the margins of these brands with those of the others would answer the question we addressed – whether these brands guarantee high margins to manufacturers. Relatively small numbers in "Market share" column in Table 1 are because of outside option as consumers did not buy any of these 7 brands 87.0% of their shopping trips. Brand 7 is a brand containing a fruit, which is thought to justify its higher retail price.

As for explanatory variables for marginal cost, we collected data of raw milk price, labor wage in four prefectures where 7 brands of yogurt are produced, international sugar price, cream price index, and international oil price.[8] Because all data were only available in monthly basis, we transformed them into weekly data by the linear filtering process employed by [15]. As for international sugar price, we multiplied it to the sugar amount each brand contains. Also, since cream is mixed in yogurt to increase fat content, we multiplied cream price index to the fat amount each brand contains. We used raw milk price as they were, and we took log for labor wage and for international oil price because their scales were of different orders of magnitude. In addition, we employed manufacturer dummy variables to incorporate firm-specific cost structure.

5 Empirical Results

5.1 Demand-Side Results

We estimate the latent class model by increasing the number of segments until there is no improvement in AIC. We find that the model with six segments maximizes AIC.[9] The parameter estimates of the optimal model with standard errors are presented in Table 2. All parameters are significant at 1% level.

[8]The information sources are as follows: Raw milk price and cream price index are obtained from the database of "Japan Dairy Association"; labor wage in four prefectures are obtained from statistical departments of corresponding prefectures; international sugar price is obtained from the database of "Agriculture & Livestock Industries Corporation"; international oil price is obtained from "U.S. Energy Information Administration."

[9]Additionally, we constructed and estimated two other models, which are a multinomial logit model without state dependence and the model with lagged brand choice dummy variable with the same number of segments to compare the fits. The suggested model achieved the lowest AIC, and we preset only the result of this model in the followings.

Table 2: Parameter estimates of the optimal model.

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6
Brand 1	1.803	-2.733	-3.299	1.635	-1.071	-4.097
	(0.0002)	(0.0003)	(0.0000)	(0.0000)	(0.0001)	(0.0002)
Brand 2	3.067	-10.24	0.119	3.973	2.577	-20.62
	(0.0001)	(0.0000)	(0.0002)	(0.0000)	(0.0021)	(0.0000)
Brand 3	2.358	0.829	-4.435	1.887	1.629	-2.142
	(0.0002)	(0.0028)	(0.0000)	(0.0000)	(0.0005)	(0.0004)
Brand 4	2.245	-1.827	3.938	-0.403	1.054	-2.943
	(0.0000)	(0.0001)	(0.0037)	(0.0000)	(0.0004)	(0.0002)
Brand 5	14.00	3.580	4.542	8.090	3.525	5.026
	(0.0002)	(0.0093)	(0.0000)	(0.0000)	(0.0000)	(0.003)
Brand 6	12.88	-0.598	-2.733	14.50	4.882	4.045
	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0002)
Brand 7	7.178	-1.093	2.702	7.478	-2.304	-1.478
	(0.0001)	(0.0001)	(0.0000)	(0.0002)	(0.0000)	(0.0000)
Price Coefficient	-17.10	-1.754	-14.30	-21.12	-10.90	-5.812
	(0.0025)	(0.0048)	(0.0042)	(0.0006)	(0.0037)	(0.0077)
Segment sizes	41.3%	2.7%	8.9%	30.4%	6.9%	9.8%
Demographics						
Intercept	0.518	4.119	1.151	-7.100	-1.525	0.099
	(0.0131)	(0.0134)	(0.0585)	(0.008)	(0.0236)	(0.0062)
Male dummy	0.696	4.138	-1.953	-1.786	0.517	-1.419
	(0.0015)	(0.0003)	(0.0001)	(0.0001)	(0.0002)	(0.0001)
Age (logged)	0.143	-2.644	1.641	2.283	0.962	-0.152
	(0.0471)	(0.0536)	(0.2405)	(0.0414)	(0.0943)	(0.0231)
The attribute similarity index						
Raw milk usage	0.060					
	(0.0016)					
Ager usage	0.358					
	(0.0012)					
Number of parameters	72					
Number of observations	15,194					
Log-likelihood	-6,666.2					

In Table 2, "Brand" entries represent brand-specific intercepts relative to outside options, presented below "Demographics" entry are parameters for calculating SD_s, which is a parameter of the attribute similarity index in (2), and presented below "The attribute similarity index" entry are the estimates of importance weights for two attributes to calculate the attribute similarity index.[10] Because we find that using all three attributes results in anomalies in estimation, we choose to remove "Fat level" attribute. The larger number of "Ager usage" relative to "Raw milk usage" suggests that perceived similarity between brands largely depends on the type of yogurt (i.e., whether yogurt is hard-type or plain-type).

We calculated how state dependence tendency varies across segments by genders using mean age. Households in segment 4 and 6 are found to be variety-seekers and the rest is almost all inertial. Only segment 2 had opposing signs for state dependence tendency depending on gender (males in this segment exhibit strong inertial tendency while females exhibit modest variety-seeking tendency). Overall, we do not see the consistent relationship between state dependence tendencies and demographic variables. Being male affects the utility of the similar brand to previously purchased one either positively or negatively, and the same is true for age.

[10]We only present estimates of importance weights for segment 1 in Table 2. This is because we estimated them with the model without segment and used these estimates for the models with the greater number of segments. In other words, we assumed perceptions of similarity between brands were common across segments as in [4]. This is because estimating the model without this assumption would have increased the number of parameters by 66, and this could have made the estimation unstable.

Table 3: Margins (Unit: yen per gram) under each model and game.

	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6	Brand 7
Average Prices	0.451	0.504	0.513	0.480	1.127	1.128	0.859
Retail margin	0.051	0.109	0.105	0.104	0.313	0.187	0.103
	(0.0006)	(0.0013)	(0.0031)	(0.0024)	(0.0051)	(0.0026)	(0.0019)
manufacturer Stackelberg							
Bertrand competition	0.039	0.100	0.007	0.125	0.039	0.115	0.085
	(0.0006)	(0.0028)	(0.0044)	(0.0015)	(0.0033)	(0.0014)	(0.0010)
Tacit collusion	0.048	0.131	0.018	0.147	0.045	0.119	0.107
	(0.0010)	(0.0023)	(0.0046)	(0.0015)	(0.0035)	(0.0013)	(0.0021)
vertical Nash							
Bertrand competition	0.042	0.084	0.088	0.084	0.304	0.182	0.083
	(0.0005)	(0.0002)	(0.0026)	(0.0016)	(0.0050)	(0.0027)	(0.0011)

5.2 Supply-Side Results

In this subsection, we will present the results of margins, marginal cost and model comparison. Though the actual calculations proceed in this order, we first present the result of model comparison as it helps the interpretation of the results of margins.

5.2.1 Log-likelihood for supply-side and Vuong test statistics

After calculating margins, we calculated the log-likelihood for supply-side in (11) and Vuong test statistics to compare the fits of three models and games in these models. We find that the market is best described by the vertical Nash–Bertrand competition game. In addition to forward-looking model, where firms account for the impact of current price on future profit, we also conducted analysis using static model and myopic model. The static model is a standard multinomial logit model without state dependence and the myopic model assumes that firms account for state dependence in demand (i.e., firms consider the effect of a household previous brand choice via the attribute similarity index) but do not account for the future profit associated with current pricing decision. We compare the Vuong test statistics across models to find that the best-fitting model (the vertical Nash–Bertrand competition game in forward-looking model) is statistically better than any other models and games.

5.2.2 Margins

The margins (in yen per gram) of suggested model are presented in Table 3. It should be noted that margins in the vertical Nash–Bertrand competition game in Table 3 are our best estimate within the employed framework, and those in the other entries are counter-factual in the sense that, had these sorts of games and perspectives were in play, these margins would have resulted. Now we briefly overview the results in Table 3.

First of all, manufacturers' margins under tacit collusion always exceed those under Bertrand competition as expected. However, for brand 1, 3, 5 and 6, the margins under manufacturer Stackelberg are lower than vertical Nash counterparts in both myopic and forward-looking models regardless of which game in horizontal interaction is assumed. This is one piece of evidence that manufacturer Stackelberg game between manufacturers and a retailer cannot be justified with data.

Remember that brand 3 has a distinct taste advantage due to the fact that it uses only raw milk, while brand 5 and 6 are the yogurt with special lactic-acid bacilli. Therefore we expect that these brands to command higher margins. As expected, brand 3, 5 and 6 command three largest margins under the vertical Nash–Bertrand competition game (0.088, 0.304, and 0.182 respectively), which we estimate to reflect Japanese yogurt market. Meanwhile, brand 3 and 5 have the least and the second least margins respectively under the manufacturer Stackelberg–Bertrand competition counter-factual (0.007 and 0.039), which is another evidence that manufacturer Stackelberg game cannot be justified with data.[11]

These facts and the market being characterized by the vertical Nash–Bertrand competition game jointly imply that differentiating brands by improving its quality enables manufacturers to charge higher margins relative to the others. However, we note that a retailer also charges the largest and the second largest margins for brand 5 and 6 and charges the fourth largest margin for brand 3. In fact, the amount of retailer's margins are higher than manufacturers' margins for all brands in the vertical Nash–Bertrand competition game as shown in Table 3. These facts lead us to the conclusion that a retailer has more power than manufacturers. The decreasing price of yogurt over the last decade is at least partially due to decreasing power of manufacturers relative to the retailer in addition to competition among manufacturers as indicated by our result. The existence of fierce competition among manufacturers makes sense, as 157 yogurt brands existed in the market in January 2007 to December 2008.

5.2.3 Marginal cost

The estimation result for marginal cost of the vertical Nash–Bertrand competition game of the proposed model is presented in Table 4.[12] We find that after including manufacturer dummy variables, all variables except for international oil price have negative coefficients in the best-fitting game, thus we exclude them.[13] The high values for manufacturers

[11] The margins of Brand 1 and 5 under the manufacturer Stackelberg–Bertrand competition game in forward-looking model appear to be the same in Table 3, but this is because of rounding. The margin of brand 1 is slightly larger than that of brand 5, even though the difference is minimal.

[12] Results for the other models can be provided upon the request to the author.

[13] If we use only labor wage, their coefficients are positive. The effect of labor wage seems to be absorbed by manufacturer dummy variables.

Table 4: Marginal cost estimation in forward-looking model.

manufacturer Stackelberg	Bertrand	competition			Tacit	collusion		
	Estimate	Std.Err	t-value	p-value	Estimate	Std.Err	t-value	p-value
Intercept	-0.742	0.233	-3.182	0.002	-0.816	0.240	-3.408	0.001
Manufacturer 2	-0.041	0.026	-1.598	0.110	-0.061	0.027	-2.319	0.021
Manufacturer 3	-0.026	0.030	-0.880	0.379	-0.031	0.031	-1.004	0.316
Manufacturer 4	0.287	0.022	13.22	0.000	0.287	0.022	12.87	0.000
Manufacturer 5	0.346	0.027	13.04	0.000	0.335	0.027	12.28	0.000
Cream price index	0.133	0.032	4.137	0.000	0.139	0.033	4.192	0.000
International oil price	0.038	0.024	1.546	0.123	0.037	0.025	1.480	0.139
Raw milk price	0.003	0.002	1.104	0.270	0.003	0.002	1.278	0.202
vertical Nash	Bertrand	competition			Tacit	collusion		
	Estimate	Std.Err	t-value	p-value	Estimate	Std.Err	t-value	p-value
Intercept	0.196	0.082	2.401	0.017	0.187	0.084	2.225	0.026
Manufacturer 2	-0.048	0.021	-2.280	0.023	-0.063	0.021	-2.918	0.004
Manufacturer 3	-0.037	0.021	-1.778	0.076	-0.044	0.021	-2.075	0.038
Manufacturer 4	0.163	0.017	9.563	0.000	0.161	0.018	9.193	0.000
Manufacturer 5	0.315	0.021	15.11	0.000	0.304	0.021	14.17	0.000
International oil price	0.037	0.018	2.015	0.044	0.037	0.019	1.957	0.051

4 and 5 are consistent with the fact that manufacturer 4 produces brand 5 and 6 and manufacturer 5 produces brand 7.

5.2.4 The price endogeneity

After estimating $\widehat{\xi_{jt}}$ and $\widehat{\eta_{jt}}$, we tested the correlation between them using one of Pearson's product moment correlation coefficient test. The test reveals that they are significantly correlated and thus prices are proven to be endogenously determined, which is consistent with the general finding in literature.

6 Conclusion

In this paper, we empirically analyzed Japanese yogurt market incorporating consumer heterogeneity, consumer state dependence, forward-looking behavior of manufacturers and a retailer, and price endogeneity arises from the interaction between unobserved demand characteristics and prices. Our demand-side findings are consistent with those of previous literature; consumers are heterogeneous in their responsiveness to marketing variables and degrees of state dependence. On supply-side, we find prices are endogenously determined, manufacturers engage in Bertrand competition game, manufacturers and a retailer play vertical Nash game, and they set prices considering their impact on future profit.

We find that brands with differentiating features (brand 3, 5 and 6) do command higher margins, proving that manufacturers' efforts are rewarded. However, a retailer also charges higher margins for these brands and obtains larger split of the profit. We also find that there are rigorous competitions among manufacturers in this market which is consistent with the findings in the other papers such as [14] and [4], where Bertrand competition was the case in the U.S. cereal market with large number of brands. Finally, our work adds another evidence to the body of literature in this field of intersection between marketing and neo empirical industrial organization, as lack of empirical study is general concern in this area [10].

One major limitation of this research is the assumption of a monopolistic retailer as retailers are likely to compete in reality. In fact, "National Survey of Prices" conducted by Statistics Bureau, Ministry of Internal Affairs and Communications in Japan indicates that the average retail prices of yogurt are higher in stores with no competitors around. Incorporating retail competition in the framework employed in this study would be an interesting source of future research. The other possible direction of future research is inclusion the effect of store brand. This topic is common in the literature, and widely investigated in the context such as its effect on power balance between manufacturers, store loyalty and so forth. As state dependence is often neglected in these analysis, investigating the effect of store brand in the presented framework may provide new insight to the literature.

Appendix

In appendix A, we derive margins in myopic model. In appendix B, we derive margins in forward-looking model. We note that equations to derive margins in static model are identical to those in myopic model.

A Margins in Myopic Model

We start with margins of a retailer as it will be used in calculating margins of manufacturers.

A.1 Margins of a Retailer

The profit function for the monopolistic retailer is defined as

$$\pi_R = \sum_{j=1}^{J}(p_{jt} - w_{jt})S_{jt}M \tag{12}$$

where w_{jt} is the wholesale price for brand j at time t, S_{jt} is the market share, and M is the market size. The retail margin for brand j is $p_{jt} - w_{jt}$.

Now by partially differentiating (12) with respect to each retail price p_{jt}, setting them zero, and algebraic manipulations, we have

$$\begin{pmatrix} p_{1t} - w_{1t} \\ \vdots \\ p_{Jt} - w_{Jt} \end{pmatrix} = - \begin{bmatrix} \frac{\partial S_{1t}}{\partial p_{1t}}, \cdots, \frac{\partial S_{Jt}}{\partial p_{1t}} \\ \vdots \\ \frac{\partial S_{1t}}{\partial p_{Jt}}, \cdots, \frac{\partial S_{Jt}}{\partial p_{Jt}} \end{bmatrix}^{-1} \begin{pmatrix} S_{1t} \\ \vdots \\ S_{Jt} \end{pmatrix}. \tag{13}$$

Using the notation of [4], we have

$$(\mathbf{p}_t - \mathbf{w}_t) = \Phi_t^{-1}\mathbf{S}_t \tag{14}$$

where $(\mathbf{p}_t - \mathbf{w}_t) \equiv (p_{1t} - w_{1t}, \cdots, p_{Jt} - w_{Jt})^T$ is $J \times 1$ vector of retail margins, Φ_t is $J \times J$ matrix with elements

$$\Phi_{jkt} = -\frac{\partial S_{kt}}{\partial p_{jt}}$$

for brand $j, k = 1, \cdots, J$, and $\mathbf{S}_t$ is $J \times 1$ vector $\mathbf{S}_t = (S_{1t}, \ldots, S_{Jt})^T$.

A.2 Margins of Manufacturers

Now we derive margins of manufacturers under different games. Unlike in the retailer's case, the profit function of manufacturers differs depending on which game in horizontal strategic interaction is assumed. The profit function π_f of manufacturer f under Bertrand competition is given by

$$\pi_f = \sum_{j \in J_f} (w_{jt} - mc_{jt}) S_{jt} M, \tag{15}$$

where J_f is a subset of brands produced by manufacturer f and mc_{jt} is the marginal cost of producing brand j at time t. The manufacturer's margin from brand j is $w_{jt} - mc_{jt}$. On the other hand, the total profit function $\pi_{\forall f}$ of collusive manufacturers is given by

$$\pi_{\forall f} = \sum_{j=1}^{J} (w_{jt} - mc_{jt}) S_{jt} M.$$

The first order condition of the profit function in tacit collusion game is

$$\frac{\partial \pi_{\forall f}}{\partial w_{lt}} = M \left[S_{lt} + \sum_{j=1}^{J} \left[(w_{jt} - mc_{jt}) \sum_{k=1}^{J} \frac{\partial S_{jt}}{\partial p_{kt}} \cdot \frac{\partial p_{kt}}{\partial w_{lt}} \right] \right] = 0 \tag{16}$$

for $l = 1, \ldots, J$. By algebraic manipulation, we have

$$\begin{pmatrix} w_{1t} - mc_{1t} \\ \vdots \\ w_{Jt} - mc_{Jt} \end{pmatrix} = - \left[\begin{bmatrix} \frac{\partial p_{1t}}{\partial w_{1t}}, \cdots, \frac{\partial p_{Jt}}{\partial w_{1t}} \\ \vdots \\ \frac{\partial p_{1t}}{\partial w_{Jt}}, \cdots, \frac{\partial p_{Jt}}{\partial w_{Jt}} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial S_{1t}}{\partial p_{1t}}, \cdots, \frac{\partial S_{Jt}}{\partial p_{1t}} \\ \vdots \\ \frac{\partial S_{1t}}{\partial p_{Jt}}, \cdots, \frac{\partial S_{Jt}}{\partial p_{Jt}} \end{bmatrix} \right]^{-1} \begin{pmatrix} S_{1t} \\ \vdots \\ S_{Jt} \end{pmatrix}, \tag{17}$$

where the left hand side of equation (17) is $J \times 1$ vector of manufacturers' margins. The first order condition of profit function in Bertrand competition can be derived similarly. In equation (17), the terms S_{jt} and $\partial S_{jt}/\partial p_{kt}$ can be directly obtained from the estimated demand parameters but $\partial p_{kt}/\partial w_{lt}$ cannot be. Thus we must infer these terms indirectly, and the difference between manufacturer Stackelberg and vertical Nash stems from how these terms are specified. We start with the manufacturer Stackelberg–Tacit collusion game because margins under the other games can be derived as the special case of this game.

A.2.1 Margins under the manufacturer Stackelberg–Tacit collusion game

To infer $\partial p_{kt}/\partial w_{lt}$, we exploit the first order condition of the retail profit function defined in (12);

$$\frac{\partial \pi_R}{\partial p_{gt}} = S_{gt} + \sum_{k=1}^{J} \left[(p_{kt} - w_{kt}) \frac{\partial S_{kt}}{\partial p_{gt}} \right] = 0 \tag{18}$$

for $g = 1, \ldots, J$ with the market size M removed. Since a retailer is assumed to maximize the category profit, the change in wholesale price of one brand would affect all retail prices in the category. Thus we totally differentiate (18) with respect to prices p_{jt}, $j = 1, \ldots, J$, and wholesale price w_{lt} for brand l, to obtain, for some g,

$$\sum_{j=1}^{J} \left[\frac{\partial S_{gt}}{\partial p_{jt}} + \frac{\partial S_{jt}}{\partial p_{gt}} + \sum_{k=1}^{J} (p_{kt} - w_{kt}) \frac{\partial^2 S_{kt}}{\partial p_{jt} \partial p_{gt}} \right] dp_{jt} - \frac{\partial S_{lt}}{\partial p_{gt}} \cdot dw_{lt} = 0. \tag{19}$$

Denoting the terms inside the bracket on the left hand side of equation (19) as $\nu(g, j)$, we have the set of J equations for some l as

$$\begin{cases} \nu(1,1) dp_{1t} + \nu(1,2) dp_{2t} + \cdots + \nu(1,J) dp_{Jt} = \frac{\partial S_{lt}}{\partial p_{1t}} \cdot dw_{lt}, \\ \vdots \\ \nu(J,1) dp_{1t} + \nu(J,2) dp_{2t} + \cdots + \nu(J,J) dp_{Jt} = \frac{\partial S_{lt}}{\partial p_{Jt}} \cdot dw_{lt}. \end{cases} \tag{20}$$

Defining $G_g \equiv (\nu(g,1), \ldots, \nu(g,J))$, we rewrite the expression in (20) in matrix form and rearrange it as

$$\begin{pmatrix} \partial p_{1t}/\partial w_{lt} \\ \vdots \\ \partial p_{Jt}/\partial w_{lt} \end{pmatrix} = \begin{pmatrix} G_1 \\ \vdots \\ G_J \end{pmatrix}^{-1} \cdot \begin{pmatrix} \frac{\partial S_{lt}}{\partial p_{1t}} \\ \vdots \\ \frac{\partial S_{lt}}{\partial p_{Jt}} \end{pmatrix}, \tag{21}$$

assuming the inverse of the $J \times J$ matrix $(G_1, \ldots, G_J)^T$ exists. Transposing both sides of equation (21) and stacking them vertically for $l = 1, \cdots, J$, we have

$$\begin{bmatrix} \frac{\partial p_{1t}}{\partial w_{1t}}, \cdots, \frac{\partial p_{Jt}}{\partial w_{1t}} \\ \vdots \\ \frac{\partial p_{1t}}{\partial w_{Jt}}, \cdots, \frac{\partial p_{Jt}}{\partial w_{Jt}} \end{bmatrix} = \begin{bmatrix} \frac{\partial S_{1t}}{\partial p_{1t}}, \cdots, \frac{\partial S_{1t}}{\partial p_{Jt}} \\ \vdots \\ \frac{\partial S_{Jt}}{\partial p_{1t}}, \cdots, \frac{\partial S_{Jt}}{\partial p_{Jt}} \end{bmatrix} \cdot \left(G_1^T, \cdots, G_J^T\right)^{-1}. \tag{22}$$

Substituting (22) into (17), we have the manufacturers' margins under the manufacturer Stackelberg–Tacit collusion game as

$$(\mathbf{w}_t - \mathbf{mc}_t) \quad = \quad -(\Phi_t^T \mathbf{G}^{-1} \Phi_t)^{-1} \mathbf{S}_t \tag{23}$$

where $(\mathbf{w}_t - \mathbf{mc}_t) = (w_{1t} - mc_{1t}, \cdots, w_{Jt} - mc_{Jt})^T$ and $\mathbf{G} = (G_1^T, \cdots, G_J^T)$.

A.2.2 Margins under the manufacturer Stackelberg–Bertrand competition game

In Bertrand competition, each manufacturer maximizes the profit from its own brands. Thus in Bertrand competition, (17) applies only to the brands a particular manufacturer produces. This requires replacement of the third term Φ_t in matrix $(\Phi_t^T \mathbf{G}^{-1} \Phi_t)^{-1}$ in (23) with $\Phi_t \cdot *\Omega$, where $\cdot *$ denotes element-by-element multiplication, and Ω is $J \times J$ matrix whose (j, k) elements are indicator functions taking unity if brands j and k are made by the same manufacturer and zero otherwise. Then we have the manufacturers' margins under the manufacturer Stackelberg–Bertrand competition game as

$$(\mathbf{w}_t - \mathbf{mc}_t) = -(\Phi_t^T \mathbf{G}^{-1} \Phi_t \cdot *\Omega)^{-1} \mathbf{S}_t. \tag{24}$$

A.2.3 Margins under the vertical Nash–Tacit collusion game

In the vertical Nash game, manufacturers and a retailer move simultaneously. More specifically, manufacturers set wholesale price expecting a certain level of retail margin for the brand; a retailer sets its retail margin for each brand based on its profit maximizing behavior. Now by assumption, we have the relationship

$$\frac{\partial (p_{jt} - w_{jt})}{\partial w_{jt}} = 0$$

or equivalently

$$\frac{\partial p_{jt}}{\partial w_{jt}} = 1 \tag{25}$$

for all $j = 1, \ldots, J$ since the retail margin of brand j, $p_{jt} - w_{jt}$, is not affected by the wholesale price of the brand as manufacturers and a retailer move simultaneously.

Similarly, since the retail margin of brand, $p_{jt} - w_{jt}$, is not affected by the wholesale price of the other brands, we have

$$\frac{\partial(p_{jt} - w_{jt})}{\partial w_{kt}} = 0$$

or equivalently

$$\frac{\partial p_{jt}}{\partial w_{kt}} = 0 \tag{26}$$

for $j = 1, \ldots, J$, $j \neq k$.[14] Finally, from (25) and (26), the matrix with elements $\partial p_{jt}/\partial w_{kt}$ on the right-hand side of equation (17) becomes an identity matrix and equation (17) becomes

$$\begin{pmatrix} w_{1t} - mc_{1t} \\ \vdots \\ w_{Jt} - mc_{Jt} \end{pmatrix} = - \begin{bmatrix} \frac{\partial S_{1t}}{\partial p_{1t}}, \ldots, \frac{\partial S_{Jt}}{\partial p_{1t}} \\ \vdots \\ \frac{\partial S_{1t}}{\partial p_{Jt}}, \ldots, \frac{\partial S_{Jt}}{\partial p_{Jt}} \end{bmatrix}^{-1} \begin{pmatrix} S_{1t} \\ \vdots \\ S_{Jt} \end{pmatrix}.$$

Thus we have manufacturers' margins under the vertical Nash–Tacit collusion game as

$$(\mathbf{w}_t - \mathbf{mc}_t) = \Phi_t^{-1}\mathbf{S}_t \tag{27}$$

which is identical to margin of the retailer. This makes sense as the vertical Nash game assumes approximately equal power between manufacturers and a retailer [6].

A.2.4 Margins under the vertical Nash–Bertrand competition game

Since the retailer behaves the same independent of whether manufacturers compete or tacitly collude, the conditions (25) and (26) still hold in the vertical Nash–Bertrand competition game. And by the same reasoning of the manufacturer Stackelberg–Bertrand competition game, we have the manufacturers' margins under the vertical Nash–Bertrand competition game as

$$(\mathbf{w}_t - \mathbf{mc}_t) = (\Phi_t \cdot *\Omega)^{-1}\mathbf{S}_t. \tag{28}$$

[14] We note that this behavioral principle of retailer is consistent with its profit maximizing behavior, as the predetermined retail margins are still determined from the first order condition of its profit function

$$\frac{\partial \pi_R}{\partial p_{gt}} = S_{gt} + \sum_{k=1}^{J}\left[(p_{kt} - w_{kt})\frac{\partial S_{kt}}{\partial p_{gt}}\right] = 0$$

even in the vertical Nash game.

B Margins in Forward-Looking Model

Here we derive the margins in forward-looking model. We start with the margin of a retailer.

B.1 Margins of a Retailer (Forward-Looking Model)

The objective function of one-period forward-looking retailer is $V_R = \pi_{R1} + \delta\pi_{R2}$, where π_{Rt} is a profit function defined in (12) for period $t = 1, 2$, and the term δ is some exogenously given discount rate. Then the first order conditions are

$$\begin{cases} \frac{\partial \pi_{R1}}{\partial p_{k1}} + \delta \sum_{j=1}^{J} \frac{\partial \pi_{R2}}{\partial S_{j2}} \cdot \frac{\partial S_{j2}}{\partial S_{j1}} \cdot \frac{\partial S_{j1}}{\partial p_{k1}} &= 0 \\ \frac{\partial \pi_{R2}}{\partial p_{k2}} &= 0 \end{cases} \tag{29}$$

for $k = 1, \ldots, J$. In (29), the first equation corresponds to the first order condition of the first period profit function and the second equation corresponds to that of the second period profit. As the first order condition in the second period is already known, we only concern for the first equation in (29) in the following derivation. Furthermore, in that equation, the unknown terms are $\partial\pi_{R2}/\partial S_{j2}$ and $\partial S_{j2}/\partial S_{j1}$.

Clearly, $\partial\pi_{R2}/\partial S_{j2}$ is $(p_{j2} - w_{j2})$. To calculate $\partial S_{j2}/\partial S_{j1}$, we exploit the following relationship:

$$S_{j2} = \theta_{j2|j1} \times S_{j1} + \sum_{l=1, l\neq j}^{J} \theta_{j2|l1} \times S_{l1} \tag{30}$$

where $\theta_{j2|j1}$ is the probability of purchasing brand j in period 2 given the purchase of the brand in period 1, and $\theta_{j2|l1}$ is defined likewise with brand l. Since the market share sums up to one, the term S_{l1} is rewritten as $S_{l1} = (1 - S_{11} - \cdots - S_{l-1,1} - S_{l+1,1} - \cdots - S_{J1})$ for all $l = 1, \ldots, J$, $l \neq j$, which includes the term $-S_{j1}$. Thus, the partial derivative of the second term on the right-hand side of equation (30) with respect to S_{j1} is

$$\frac{\partial \left[\sum_{l=1, l\neq j}^{J} \theta_{j2|l1} \times S_{l1}\right]}{\partial S_{j1}} = -\sum_{l=1, l\neq j}^{J} \theta_{j2|l1}$$

as $\partial S_{l1}/\partial S_{j1} = -1$ for $l = 1, \ldots, J$, $l \neq j$. Thus taking partial derivative of both sides of (30) with respect to S_{j1}, we have

$$\frac{\partial S_{j2}}{\partial S_{j1}} = \theta_{j2|j1} - \sum_{l=1, l\neq j}^{J} \theta_{j2|l1}. \tag{31}$$

We define the right-hand side of equation (31) as Δ_j.

In the same manner as in the derivation of vector $(\mathbf{p}_t - \mathbf{w}_t)$, the second term on the left-hand side of the first equation in (29) can be expressed by matrix form as

$$\delta \begin{bmatrix} \frac{\partial S_{11}}{\partial p_{11}}, \cdots, \frac{\partial S_{J1}}{\partial p_{11}} \\ \vdots \\ \frac{\partial S_{11}}{\partial p_{J1}}, \cdots, \frac{\partial S_{J1}}{\partial p_{J1}} \end{bmatrix} \cdot \begin{bmatrix} \Delta_1, \cdots, 0 \\ \vdots \\ 0, \cdots, \Delta_J \end{bmatrix} \cdot \begin{pmatrix} p_{12} - w_{12} \\ \vdots \\ p_{J2} - w_{J2} \end{pmatrix}$$

where the second matrix is diagonal matrix with diagonal elements Δ_j, which we will express as $\mathbf{\Delta}$. Thus we have the margin in the first period as

$$\begin{pmatrix} p_{11} - w_{11} \\ \vdots \\ p_{J1} - w_{J1} \end{pmatrix} = - \begin{bmatrix} \frac{\partial S_{11}}{\partial p_{11}}, \cdots, \frac{\partial S_{J1}}{\partial p_{11}} \\ \vdots \\ \frac{\partial S_{11}}{\partial p_{J1}}, \cdots, \frac{\partial S_{J1}}{\partial p_{J1}} \end{bmatrix}^{-1} \cdot \begin{pmatrix} S_{11} \\ \vdots \\ S_{J1} \end{pmatrix} - \delta \begin{bmatrix} \Delta_1, \cdots, 0 \\ \vdots \\ 0, \cdots, \Delta_J \end{bmatrix} \cdot \begin{pmatrix} p_{12} - w_{12} \\ \vdots \\ p_{J2} - w_{J2} \end{pmatrix}$$

or $(\mathbf{p}_1 - \mathbf{w}_1) = \{\Phi^T\}^{-1}\mathbf{S}_1 - \delta\mathbf{\Delta}(\mathbf{p}_2 - \mathbf{w}_2)$, assuming the inverse of Φ^T exists. To derive margins in forward-looking model, we first calculate the margins in the myopic case from week 2, and use these margins in calculating margins in forward-looking model starting from week 1.

B.2 Margins of Manufacturers (Forward-Looking Model)

The derivation of margins of manufacturers in one-period forward-looking model much follows the case of the retailer. Here we consider the margin in the manufacturer Stackelberg–Tacit collusion game as margins in the other games are special case of those under this game. The objective function is $V_M = \pi_{f1} + \delta\pi_{f2}$ and the first order conditions are

$$\begin{cases} \frac{\partial \pi_{f1}}{\partial w_{k1}} + \delta \sum_{j=1}^{J} \frac{\partial \pi_{f2}}{\partial S_{j2}} \cdot \frac{\partial S_{j2}}{\partial S_{j1}} \cdot \frac{\partial S_{j1}}{\partial w_{k1}} &= 0 \\ \frac{\partial \pi_{f2}}{\partial w_{k2}} &= 0. \end{cases} \tag{32}$$

As was the case in (29), the first equation of (32) corresponds to the first order condition of the first period profit function and the second equation corresponds to that of the second period profit. Clearly, $\partial\pi_{f2}/\partial S_{j2} = (w_{j2} - mc_{j2})$. Then the product of this term and $\partial S_{j1}/\partial w_{k1}$ turns out to be the second term of the first order condition of the profit function of manufacturers in (16), except for the subscript being 2 instead of t in wholesale price w_{j2} and marginal cost mc_{j2}. Then this product term can be written in matrix form as

$$\begin{bmatrix} \frac{\partial p_{11}}{\partial w_{11}}, \cdots, \frac{\partial p_{J1}}{\partial w_{11}} \\ \vdots \\ \frac{\partial p_{11}}{\partial w_{J1}}, \cdots, \frac{\partial p_{J1}}{\partial w_{J1}} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial S_{11}}{\partial p_{11}}, \cdots, \frac{\partial S_{J1}}{\partial p_{11}} \\ \vdots \\ \frac{\partial S_{11}}{\partial p_{J1}}, \cdots, \frac{\partial S_{J1}}{\partial p_{J1}} \end{bmatrix} \cdot \begin{pmatrix} w_{12} - mc_{12} \\ \vdots \\ w_{J2} - mc_{J2} \end{pmatrix}$$

or simply $\Phi_t^T \mathbf{G}^{-1}\Phi_t(\mathbf{w}_2 - \mathbf{mc}_2)$. Thus the second term on the left-hand side of the first equation of (32) becomes $\delta(\Phi_t^T \mathbf{G}^{-1}\Phi_t)\boldsymbol{\Delta}(\mathbf{w}_2 - \mathbf{mc}_2)$. Then we have $\mathbf{S}_1 + \Phi_t^T \mathbf{G}^{-1}\Phi_t(\mathbf{w}_1 - \mathbf{mc}_1) + \delta(\Phi_t^T \mathbf{G}^{-1}\Phi_t)\boldsymbol{\Delta}(\mathbf{w}_2 - \mathbf{mc}_2) = 0$ or $(\mathbf{w}_1 - \mathbf{mc}_1) = -(\Phi_t^T \mathbf{G}^{-1}\Phi_t)^{-1}\mathbf{S}_1 - \delta \cdot \boldsymbol{\Delta}(\mathbf{w}_2 - \mathbf{mc}_2)$, assuming the inverse of $\Phi_t^T \mathbf{G}^{-1}\Phi_t$ exists. The margins in the other games are derived similarly as we presented in the myopic case.

References

[1] Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

[2] David Besanko, Jean-Pierre Dubé, and Sachin Gupta. Competitive price discrimination strategies in a vertical channel using aggregate retail data. *Management Science*, 49(9):1121–1138, 2003.

[3] David Besanko, Sachin Gupta, and Dipak Jain. Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science*, 44(11-part-1):1533–1547, 1998.

[4] Hai Che, K Sudhir, and PB Seetharaman. Bounded rationality in pricing under state-dependent demand: Do firms look ahead, and if so, how far? *Journal of Marketing Research*, pages 434–449, 2007.

[5] Pradeep K Chintagunta. Investigating category pricing behavior at a retail chain. *Journal of Marketing Research*, 39(2):141–154, 2002.

[6] S Chan Choi. Price competition in a channel structure with a common retailer. *Marketing Science*, 10(4):271–296, 1991.

[7] Michaela Draganska and Dipak C Jain. Consumer preferences and product-line pricing strategies: An empirical analysis. *Marketing science*, 25(2):164–174, 2006.

[8] Michaela Draganska and Daniel Klapper. Retail environment and manufacturer competitive intensity. *Journal of Retailing*, 83(2):183–198, 2007.

[9] Dan Horsky and Paul Nelson. New brand positioning and pricing in an oligopolistic market. *Marketing Science*, 11(2):133–153, 1992.

[10] Vrinda Kadiyali, K Sudhir, and Vithala R Rao. Structural analysis of competitive behavior: New empirical industrial organization methods in marketing. *International Journal of Research in Marketing*, 18(1):161–186, 2001.

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

[12] James M. Lattin. A model of balanced choice behavior. *Marketing Science*, 6(1):48–65., 1987.

[13] Sergio Meza and K Sudhir. Do private labels increase retailer bargaining power? *Quantitative Marketing and Economics*, 8(3):333–363, 2010.

[14] Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342, 2001.

[15] Margaret E Slade. Product rivalry with multiple strategic weapons: An analysis of price and advertising competition. *Journal of Economics & Management Strategy*, 4(3):445–476, 1995.

[16] S Sriram and Vrinda Kadiyali. Empirical investigation of channel reactions to brand introductions. *International Journal of Research in Marketing*, 26(4):345–355, 2009.

[17] K Sudhir. Structural analysis of manufacturer pricing in the presence of a strategic retailer. *Marketing Science*, 20(3):244–264, 2001.

[18] J Miguel Villas-Boas and Russell S Winer. Endogeneity in brand choice models. *Management Science*, 45(10):1324–1338, 1999.

[19] J Miguel Villas-Boas and Ying Zhao. Retailer, manufacturers, and individual consumers: Modeling the supply side in the ketchup marketplace. *Journal of Marketing Research*, pages 83–95, 2005.

[20] Sha Yang, Yuxin Chen, and Greg M Allenby. Bayesian analysis of simultaneous demand and supply. *Quantitative Marketing and Economics*, 1(3):251–275, 2003.

Graduate School of Systems and Information
Division of Policy and Planning Sciences
University of Tsukuba
Tsukuba 305-8573
JAPAN
E-mail address: `tkamai@sk.tsukuba.ac.jp`
E-mail address: `kanazawa@sk.tsukuba.ac.jp`

筑波大学システム情報工学研究科　構井　友人
筑波大学システム情報系社会工学域　金澤　雄一郎