ASIA MARKETING JOURNAL

Volume 21 | Issue 3 Article 3

10-31-2019

# Competitive Analysis among Multi-product Firms

Jun B. Kim

http://dx.doi.org/10.15830/amj.2019.21.3.47

# Competitive Analysis among Multi-product Firms*

Jun B. Kim**

We analyze and study competition in differentiated product market using public data source. Understanding competitive market structure is critical for firms to assess how their products compete against other firms in a given market. In this paper, we estimate consumer demand, extend clout and vulnerability framework, and study competition among multi-product manufacturers in differentiated product market. For our empirical analysis, we adopt choice-based aggregate demand model and estimate consumer demand while accounting for unobserved product characteristics. Once we estimate consumer demand, we compute full price elasticity matrix and investigate intra- and inter- manufacturer substitutions among consumers. This research offers a framework for marketers to analyze and understand market structures, leading them to informed decisions.

Key words: Choice model, Aggregate demand model, Competitive analysis, Intra- and Inter-manufacturer competition

## Ⅰ. Introduction

Understanding consumer demand and the underlying market structure is critical to any manufacturer. Doing so, they gain insights on market participants such as consumers and competitors, leading them to informed business decisions. Manufacturers can estimate how much consumers value their products and how they will react to changes in products and prices. In addition, they can assess how other firms compete in the market. With their critical roles in mind, this paper has two main objectives. First, we aim to offer an approach to estimate consumer demand in differentiated products market while accounting for unobserved product characteristics. Second, we investigate intra- and inter- competition among multi-product manufacturers. We achieve both objectives

* This study was supported by the Institute of Management Research at Seoul National University.
** Associate Professor at Seoul National University (junbkim@snu.ac.kr)

using public data and an empirical framework that is suited for such a data set. Our target audience is managers in differentiated product industries who would like to gain insights on their markets.

Since the seminal paper by Guadagni and Little (1983), the wide availability of scanner panel data has allowed marketers to achieve a detailed understanding of consumer behaviors in most consumer-packaged goods (CPG) categories. In scanner panel data, researchers can directly observe the purchase and switching behaviors of consumers over an extended period of time, which is critical to analyze consumer preferences in the categories. However, the same does not apply to differentiated products such as consumer electronics and automobiles due to their long inter-purchase time.[1] This means that it is very challenging for researchers to observe the repeat purchase behaviors of consumers over a reasonable time window. Instead, aggregate level data such as sales or market share data are often available in differentiated products. For instance, Berry et al. (1995) used annual automobile sales data for 20 years, developed a choice-based aggregate demand model while accounting for optimal pricing setting behaviors of manufacturers, and studied consumer demand in automobile industry. Since then, some marketing academics have adopted similar approaches and studied consumer demand in diverse product categories (e.g., Sudhir 2001; Bruno 2008).[2]

However, data landscape has greatly changed since then. With the ever decreasing IT hardware cost, an unprecedented amount of online consumer data of all varieties are being generated at an unprecedented speed, leading to an era of "big data" (Erevelles et al. 2016). This includes rich data set on differentiated products[3] and Amazon.com's rich data set is a prime example. Amazon.com currently counts about two thirds of Americans as its customers and accounts for almost 50% of e-commerce in US market.[4] On the supply side, Amazon now has more than 2.5 million sellers, offering 606 million products on its US online store.[5] Its most popular product category is consumer electronics and the gross sales in this category have eclipsed those of Bestbuy in 2018, the largest US offline electronics retailer (Howland 2018).

On the data front, Amazon is the big three of "big data": Amazon, Google, and Facebook

1) According to NPD report in 2018, an average US consumer is reported to replace her smart phone in about 32 months.
2) For a comprehensive review on this topic, please refer to Kadiyali et al. (2001).
3) For marketing's perspective on this topic, please refer to Wedel and Kannan (2016).
4) Please see the link at emarketer.com (https://www.emarketer.com/content/digital-investments-pay-off-for-walmart-in-ecommerce-race)
5) Please see the link at scrapehero.com (https://www.scrapehero.com/how-many-products-does-amazon-sell-worldwide-october-2017/)

(Wedel and Kannan 2016). Among the three, Amazon.com collects and aggregates consumer transactions data and publishes the resulting sales rank data in all product categories. In addition to the sales rank data, it also makes available very detailed data on product characteristics. Recognizing the commercial potential of Amazon.com's public data, there is a growing industry that offers services on Amazon's platform. For instance, vendors such as junglescout.com, sellics.com, and amalyze.com offer to Amazon sellers sales assistance as well as marketing research based on Amazon.com's public data. Among them, sellics.com offers a predictive service that estimates the sales popularity for a product idea, which is based on Amazon's public data. Marketing academics in the past also used Amazon's data to investigate important topics in marketing (e.g., Chevalier and Mayzlin 2006; Chong et al. 2017).

Amazon.com's public data have both strengths and weaknesses compared to the data used in the past to study consumer durable goods categories. First, given its large customer base and product assortment, Amazon data may very well summarize and represent consumers' purchase behaviors in a vast array of product categories. In addition, it publishes sales rank data on a frequent basis (e.g., on a daily basis) compared to annual or quarterly data used in the past. Last, all of its data are free to use compared to high cost from 3rd party marketing research companies such as Nielsen and NPD. On the other hand, the weakness of Amazon data is their coarseness since Amazon.com offers the data in the form of sales rank. This contrasts to the sales quantity or market share data that were used in the past. Therefore, in this paper we propose a demand estimation approach that leverages the strength and accommodates the weakness of Amazon's sales rank data. Utilizing the framework, we study the intra- and inter-competition among multi-product manufacturers.

More and more consumers are choosing online channels over offline ones. And many online retail platforms such as Bestbuy.com and eBay.com collect and process consumer transactions and offer data sets similar to Amazon.com. With the right set of tools, these data may open opportunities for firms to analyze and better understand their markets. That is, they can estimate consumer demand and gain insights on their customers and competitors. This paper aims to offer such a tool and demonstrate its value in the empirical application.

This paper is organized as follows. Next section reviews the related literature, followed by the description of data collected from Amazon.com. After discussing the empirical model and its estimation, we investigate intra- and inter- competition among multi-product manufacturers. We then conclude the paper.

## Ⅱ. Related Literature

In this section, we briefly review two research streams that are related to this paper. First, we review choice-based aggregate demand models with an attention to the estimation of unobserved product characteristics. Next, we review marketing literature on techniques on visualizing market structures among competing brands.

Since the seminal paper of Berry et al. (1995), choice-based aggregate demand models have become very popular, often serving as a workhorse for many demand studies (Knittel et al. 2014). The key advantage of choice-based framework is the model parsimony. For instance, consider an empirical setting in which we are interested in estimating the full price elasticity matrix with $J$ differentiated products. If one were to adopt a linear model while fully accounting for asymmetric competition, one would consider a system of linear regression equations. In the equation, $s_{jt}$, product $j$' sales at time $t$ is expressed as,

$$s_{jt} = \alpha_j + \alpha_t + \sum_{i=1}^{J} \beta_{ji} \cdot p_{it} + \epsilon_{jt},$$

where $i$, $j=1,\cdots, J$, and $t=1,\cdots, T$. $\alpha_j$ and $\alpha_t$ are product- and time-fixed effects, and $p_{it}$ is $i$'s price at $t$. Coefficients of $\beta_{ji}$ are price coefficients among $J$ options, in which $\beta_{ji}$ is the sensitivity of $j$'s sales on $i$'s price. $\epsilon_{jt}$ are i.i.d. error term across $j$ and $t$. From this set of equations, we need to estimate $J^2$ price coefficients of $\beta_{ji}$. When $J=131$ (which is our empirical context), we need to estimate over 17,000 parameters, which is extremely high. Even if one is willing to assume symmetric price responses, i.e., $\beta_{ij} = \beta_{ji}$, the number of parameters to estimate will be over 8,500. Estimating this high number of parameters will be very demanding on the data and hence their estimation will be very challenging in many empirical contexts. Lancasterian approach dramatically reduces the number of model parameters by adopting the view that products can be characterized as a bundle of attributes (Lancaster 1966). For instance, in our empirical context of camcorders, a product will be characterized as a bundle of attributes such as brand, media format, and price among others. Once you estimate consumer preferences on these product characteristics, you can simulate and estimate cross price elasticity matrix. However, one disadvantage of the Lancasterian approach is that it fails to capture any unobserved characteristics (to analysts) such as designs, ergonomics, and sometimes, advertising when modeling a product. There are two main approaches in choice-based demand models to address this issue: contraction mapping (Berry et al. 1995) and control function (Petrin et al. 2010). However, they are not directly applicable or very challenging to apply to our empirical setting for the following reasons. While contraction mapping requires continuous variable such as sales quantity or market share as dependent variables, we have sales rank data, a set of discrete values, as our dependent variables. This makes the inversion process infeasible during the contraction mapping process. Next, although both approaches require a ready availability of instrumental variables (IV) for price, the identification of instrumental variable is quite challenging in our empirical context since we have data from one store. Therefore, we do not have empirical opportunities similar to Nevo (2001). In addition, strong instruments are often quite difficult to find in practice (Stock et al. 2002) and a weak instrument will lead to a biased estimation of model parameters (Bound et al. 1995; Stock et al. 2002). The last approach to address the unobservable product characteristics is rather simple: one can include dummy variables in consumer utility to capture product heterogeneity (Nevo 2006). This approach is feasible in our empirical setting since we observe the sales performance of the same products over time.[6] Note that this was not the case for differentiated products in the past. For instance, Berry et al. (1995) observe annual sales of over 100 nameplates over 20 years in car category. Since manufacturers typically refresh their models every year, researchers do not observe the same products over years. In contrast, we leverage the longitudinal nature of Amazon.com's data, introduce dummy variables in consumer utility, and directly account for unobserved product characteristics.

Next, we review the techniques that visualize competition and market structure with a special attention to marketing applications. To that end, we confine our discussion on multidimensional scaling (MDS) and clout and vulnerability chart (hereafter CV chart). MDS is a set of statistical methods for uncovering the relative positions of objects in a latent space by exploring their similarities or dissimilarities. In the map, the Euclidean distance among the objects is interpreted as the level of competition. That is, the brands that are located close to each other implies a high level of competition while brands that are far apart face less competition. For a comprehensive review on this topic in marketing context, please refer to Carroll and Greene (1997). MDS had found many applications in marketing (Katahira, 1990; MacKay et al., 1986; DeSarbo and Rao, 1986). Although popular among practitioners for its intuitiveness and simplicity, MDS is subject to a few disadvantages and its key disadvantage is that it is symmetric and cannot represent the asymmetric competition among

6) In practice, you need at least two observations of same product.

brands. DeSarbo, Grewal, and Wind (2006) therefore proposed a stochastic MDS model and analyzed the asymmetric competitive market structure in luxury automobile and portable phone markets.

Another popular method to visualize market competition is the clout and vulnerability chart by Kamakura and Russel (1989). The key premise behind this approach is that cross price elasticity matrix is informative of consumer substitution patterns among brands and hence can be used to describe a market structure (Allenby 1989). Extending this notion, they suggest a method to summarize and visualize the full price elasticity matrix in a concise manner. They define the clout of a focal brand $i$ as,

$$clout_i = \sum_{j \neq i}^{J} e_{ji}^2$$

where $j$ indexes the rest of the brands in the market and $e_{ji}$ is $j$'s elasticity with respect to $i$'s price change. Vulnerability of $i$ is defined with $i$ and $j$ switched in the above equation. Then, the clout and vulnerability of each brand are placed in a two-dimensional space. Several marketing researchers have adopted the CV chart to visualize the market structure implied by their models. Bronnenberg et al. (1996) used CV chart to visualize the local competition in a consumer package goods. Van Heerde et al. (2004) and Rutz et al. (2014) visualize the transitory market structure during the entry of a new brand in consumer packaged good category. Last, Bruno et al. (2008) used CV chart to visualize the market structure among confectionanry products. Common to the above papers are that they focus on consumer packaged goods category with a limited number of brands in their markets. Given a large number of products by multiple manufacturers in differentiated products market, products may compete against products from the same manufacturer as well as from other manufacturers. In order to address the intra- and inter- manufacturer competition in differentiated products market, we propose unfolding the conventional CV chart. Doing so, we aim to visualize richer patterns of competition among multi-product manufacturers. In the next section, we discuss our data for our empirical analysis.

# Ⅲ. Data

For our empirical analysis, we use aggregate-level, longitudinal data set in digital camcorder category from Amazon.com.[7] Data were collected, once every other day, for a duration of 10

7) Our empirical data in this paper is similar to Kim (2019).

months, starting from the first week of August 2006. During our data collection period, there are more than five major manufacturers offering over 300 products. Sony is the largest manufacturer in terms of product assortment. Time-invariant product characteristics include brands, media formats, pixel numbers, and form factor (compactness) among others. In addition, time-varying characteristics include sale price, average consumer ratings (i.e., average number of stars with 1 being the lowest and 5 the highest), and the number of consumer reviews. We apply the following filters to narrow down the set of products. We first remove products with missing values such as prices and confine our analysis to top five manufacturers. We also remove any professional grade products. This reduces the number of products to 132 for our empirical application.

Next, we aggregate daily sales ranks, prices, and consumer reviews on a weekly basis. The weekly average number of products in choice set is about 80 products with a minimum of 61 and a maximum of 103. Table 1 provides the descriptive statistics of the products in our empirical analysis. In this table, time-varying characteristics such as price and consumer reviews are averaged across products and time. Among the characteristics, "Seller (Amazon.com)" and "Seller (3rd party)" indicate that a product is available for purchase from Amazon.com and from 3rd party vendors, respectively. "Seller (Request)" indicates that a product is currently unavailable but consumers can submit a request to participating vendors for purchase.

## Ⅳ. Model

We develop our model with a keen attention to our empirical setting. Overall, our approach

〈Table 1〉 Product summary statistics

| Product Characteristics | Values |
|---|---|
| Brand | Sony (40), Panasonic (30), Canon (23), JVC (26), Samsung (13) |
| Media Formats | MiniDV (57), DVD (38), HD (27), FM (10) |
| Form Factor | Compact (11), Conventional (121) |
| High Definition | Yes (12), No (120) |
| Number of Pixels | 1.38M (1.00M) |
| Zoom | 19.10 (10.35) |
| Price | $533 ($291) |
| Seller (Request) | 0.01 (0.03) |
| Seller (Amazon) | 0.25 (0.42) |
| Seller($3^{rd}$ party) | 0.74(0.42) |
| Number of reviews | 9.28 (10.24) |
| Average consumer ratings | 3.03 (1.59) |

follows the choice models for aggregate-level data (e.g., Berry et al. 1995). Utility of consumer $i=1,\cdots,I$ for option $j=1,\cdots,J_t$ at week $t=1,\cdots,T$ is represented as,

$$u_{ijt} = [Z_j\,;\,X_{jt}]' \cdot \beta_i - p_{jt} \cdot \alpha_i + \xi_{jt} + e_{ijt} \quad (1)$$

in which $Z_j$ is a vector of time-invariant product characteristics, $X_{jt}$ is a vector of $j$'s time-varying product characteristics, and $\beta_{ii}$ is a vector of consumer-specific sensitivity for product characteristics. In addition, $p_{jt}$ is $j$'s price at $t$ and $\alpha_i$ is $i$'s price sensitivity. $\xi_{jt}$ is the unobserved product characteristics, a structural error term that is observed by consumers but not by analysts during choice. The last term of $e_{ijt}$ represents consumers' idiosyncratic taste, is a GEV type I random error term, and is assumed to be identical and independent across $i$, $j$, and $t$. We assume a normal distribution for consumer tastes,

$$\beta_i \sim N(b,\, \Sigma_b) \quad (2\text{-}1)$$

where b is a vector and $\Sigma_b$ is a diagonal variance-covariance matrix. In addition, we assume a log normal distribution for the price coefficient (Lee et al., 2006),

$$\log(\alpha_i) \sim N(b_\alpha,\, \sigma_\alpha^2). \quad (2\text{-}2)$$

Lastly, we discuss the unobserved product characteristics. As we discussed earlier, the discreteness of sales rank as our dependent variable does not allow us to apply the contraction mapping (BLP 1995) to estimate $\xi_{jt}$ in our model. Therefore, our strategy is to decompose this term and to see if we can still estimate some fraction of this quantity. Without loss of generality, we decompose $\xi_{jt}$ as follows,

$$\xi_{jt} = \xi_j + \xi_t + \triangle\xi_{jt}. \quad (3)$$

First, among the decomposed quantities, we cannot estimate $\triangle\xi_{jt}$ since the estimation of this quantity would require continuous dependent variable in contraction mapping. Second, our sales rank data do not allow us to estimate $\xi_t$. Note that the category sales level difference across time identifies $\xi_t$. For instance, if the category sales at $t$ are greater than those at $s$, due to seasonality, $\xi_t$ will be greater than $\xi_s$. However, with sales rank data, most popular products at $t$ and $s$ will both have sales rank of 1 even though their absolute sales levels may be different. Therefore, although sales rank data are informative about relative sale levels within same time window, they are not informative about absolute sales levels across time. Still, exclusion of $\xi_t$ is not limiting for the estimation of consumer preferences since it does not affect the competitive aspects of the products. It has a common effect on all products at $t$ in consumer utility.

Excluding the above two terms, we rewrite our utility function as,

$$u_{ijt} = [Z_j\,;\,X_{jt}]' \cdot \beta_i - p_{jt} \cdot \alpha_i + \xi_j + e_{ijt}. \quad (4)$$

We can now interpret $\xi_j$ as product-specific intercept. Conditional on product characteristics and model parameters, we can express $i$'s choice probability for $j$ at $t$ as,

$$p_{ijt}(\mathbf{Z}, \mathbf{X}; \theta_i, \boldsymbol{\xi}) = \frac{exp([\,Z_j\,;\,X_{jt}\,]' \cdot \beta_i - p_{jt} \cdot \alpha_i + \xi_j)}{\sum_{k=1}^{J_t} exp([\,Z_k\,;\,X_{kt}\,]' \cdot \beta_i - p_{kt} \cdot \alpha_i + \xi_k)}, \quad (5)$$

and the market share for $j$ at $t$ is computed by integrating out the choice probabilities across consumers,

$$S_{jt}(\mathbf{Z}, \mathbf{X}; \Theta, \boldsymbol{\xi}) = \int p_{ijt}(\mathbf{Z}, \mathbf{X}; \theta_i, \boldsymbol{\xi}) \cdot f(\theta_i|\Theta) d\Theta, \quad (6)$$

where $\Theta = \{b,\ \Sigma_b,\ \alpha,\ \sigma_\alpha^2\}$ is a vector of consumer preference parameters to estimate.

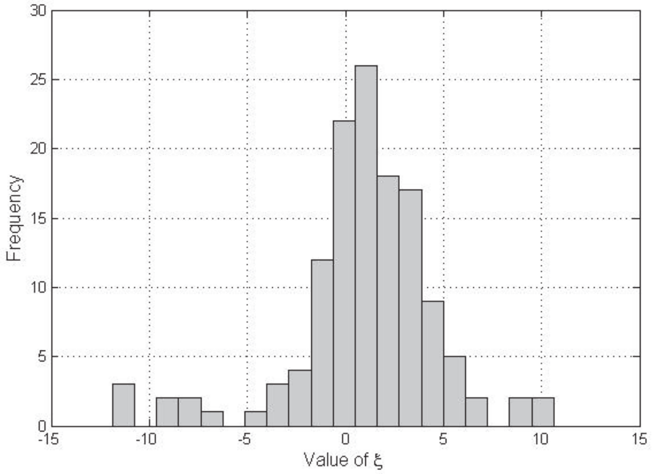
## Ⅴ. Empirical Analysis

We next describe in detail the model estimation in two separate steps. In the first step, we estimate the unobserved characteristics of $\xi_j$ outside our full model. Next, we estimate the proposed random coefficient model using $\xi_j$ estimated from the first step as additional product characteristic in consumer utility. Note that we must estimate a random coefficient choice model since a logit model suffers from IIA and does not allow flexible substitution patterns among the products. However, a flexible substitution is critical to estimate a realistic price elasticity matrix. We elaborate these two steps below.

In the first step, we use a multinomial logit with the following specification for the estimation of $\xi_j$,

$$u_{ijt} = [Z_j\,;\,X_{jt}]' \cdot \beta - p_{jt} \cdot \alpha + \xi_j^1 + e_{ijt},$$

in which all product-related vectors are defined in the same way as in Equation (1). Note that the parameters in the above equation capture mean effects of product attribute on the market outcome. Superscript 1 in $\xi_j$ means that it is a value from the first step. $\xi_j^1$ in this specification captures the remaining mean effect of $j$ conditional on the observed product characteristics. Note that the estimation of $\xi_j$ is possible due to our longitudinal data and that the estimation of multinomial logit is much faster since it does not involve high dimensional integration. Figure 1 shows the distribution of estimated $\xi_j^1$. We then treat the estimated values of $\xi_j^1$ as another product attribute in Equation (3). Given the scale difference, we substitute $\xi_j = \lambda \xi_j^1$ in Equation




〈Figure 1〉 Distribution of estimated unobservable product characteristics in the first step

(3) similar to an approach in control function (Petrin et al., 2010).

In the second step, our overall estimation strategy follows the recipe common in choice-based aggregate demand models. That is, we draw one consumer from the joint distribution of Equations (2-1) and (2-2), and predict her choice probability. We then repeat this process across different draws of consumers and aggregate their choice probabilities to predict the market share. We use identical 1,000 consumer draws at each time. For the new set of dependent variables, we closely follow the empirical approach in Kim (2019) and convert sales rank data into a set of pairwise indicator variables. This approach develops the new dependent variable of $I_{jkt}=1$ iff $r_{jt} > r_{kt}$, where $r_{jt}$ is $j$'s sales rank at time $t$. That is, if $j$ is more popular than $k$ at $t$, we set $I_{jkt}=1$ and $I_{jkt}=0$ otherwise. Once we have $\hat{S}_{jt}$ from Equation (6),we model that this quantity is associated with the unobserved, true market share of $S_{jt}$ subject to a random error term,

$$\hat{S}_{jt} = S_{jt} + \epsilon_{jt},$$

in which $\epsilon_{jt} \sim N\left(0, \frac{\sigma_\epsilon^2}{2}\right)$ is an i.i.d. random variable across $j$ and $t$. The probability of observing a pairwise rank inequality between $j$ and $k$ at $t$ is computed as,

$$\begin{aligned}\Pr(I_{jkt}=1) &= \Pr(S_{jt} > S_{kt}) \\ &= \Pr(\epsilon_{kjt} < \hat{S}_{jt} - \hat{S}_{kt}),\end{aligned}$$

where $\epsilon_{kjt} = \epsilon_{kt} - \epsilon_{jt}$ is an i.i.d. random variable with $\epsilon_{kjt} \sim N(0, \sigma_\epsilon^2)$. Therefore,

$$\Pr(I_{jkt} = 1) = \Phi\left(\frac{\hat{S}_{jt}(\mathbf{Z}, \mathbf{X}; \Theta, \boldsymbol{\xi}) - \hat{S}_{kt}(\mathbf{Z}, \mathbf{X}; \Theta, \boldsymbol{\xi})}{\sigma_\epsilon}\right),$$

where $\Phi$ is CDF of standard normal distribution and $\Theta$ are model parameters. Our likelihood function is,

$$L = \prod_{t=1}^{T}\prod_{j=1}^{J_t}\prod_{k\neq j}^{J_t}\Phi\left(\frac{\hat{S}_{jt}(\mathbf{Z}, \mathbf{X}; \Theta, \boldsymbol{\xi}) - \hat{S}_{kt}(\mathbf{Z}, \mathbf{X}; \Theta, \boldsymbol{\xi})}{\sigma_\epsilon}\right).$$

We use the Bootstrap resampling techniques for our standard error computation as in Kim (2019).

We briefly discuss the performance of our proposed model against the one without $\xi$. We confirm that our model with $\xi$ fits the data much better since AIC of our model is 81,622 while that of the model without product-intercepts is 129,200.[8] Next, as expected, the price elasticity is severely under-estimated without product-specific intercepts: its average own price elasticity is -0.56 while the proposed model estimates the value at -1.89. Therefore, they collectively confirm that it is critical to account for the unobserved product characteristics as we do in our empirical model.

〈Table 2〉 Estimated model parameters

| Model Parameters | Mean (s.e.) | Heterogeneity (s.e.) |
|---|---|---|
| Panasonic | -2.32(0.09) | 1.35(0.9) |
| Canon | -2.03(0.09) | 1.35(0.9) |
| JVC | -5.30(0.14) | 1.35(0.9) |
| Samsung | -3.34(0.17) | 1.35(0.9) |
| DVD | 0.08(0.03) | 0.32(0.2) |
| Flash Memory | 3.89(0.09) | 0.32(0.2) |
| Hard Drive | -0.83(0.05) | 0.32(0.2) |
| Compact | -2.92(0.05) | 0.06(0.04) |
| Hi-def | 1.58(0.06) | 2.32(0.04) |
| Zoom | 0.15(1e-3) | 1e-3(1e-3) |
| Pixel (in MM) | 1.34(0.02) | 0.11(0.02) |
| Xi | 0.74(0.01) | NA |
| log(Price in hundreds) | -0.88(0.02) | .23(0.1) |
| Average consumer rating | 0.13(1e-3) | .20(1e-2) |
| Number of reviews | -3e-3(1e-4) | .01(1e-3) |
| Out of stock | -0.33(0.06) | .01(1e-2) |
| Aggregation error | .01(1e-5) | |

8) The estimation result is available upon request from the authors.

The estimation result of the random coefficient model is shown in Table 2. In this table, the mean coefficient of Sony brand is normalized to 0 along with the baselines of other categorical variables. Among the continuous product characteristics, an average consumer prefers camcorders with higher pixel numbers and higher zoom. These consumer preference parameters are informative about the consumer demand. In the next section, we would like to use them to compute cross price elasticity matrix and study competition among multi-product manufacturers.

## Ⅵ. Intra- and Inter-Manufacturer Competition

We study competition among multi-product manufacturers in camcorder category using the demand parameters estimated in the previous section. After that, we summarize and visualize consumers' intra- and inter-manufacturer substitution patterns. Intra-manufacturer substitution means that consumers switch to products within same manufacturer when the focal manufacturer raises the price of one of its products. Inter-manufacturer substitution means the opposite: consumers switch away to other manufacturers' products when the manufacturer raises the price of one of its products. Therefore, a manufacturer with a high intra- and low inter-substitution means that a large fraction of its customers will still stay with the focal manufacturer and less will switch away to other manufacturers' products. The firm with low intra- and high inter- substitution means that its customers will switch away to other manufacturers. This notion is important since it has an important implication on pricing. This is so since cross price elasticity is related with the closeness of substitutes and the extent of substitutability places constraints on prices (Hausman et al. 1994). That is, if a focal manufacturer has a large fraction of its consumers substituting within a manufacturer's products, it faces less constraint on its product line pricing and can maintain elevated price levels. On the other hand, if a focal firm is subject to a high level of inter-manufacturer substitution, its pricing decision is constrained by other manufacturers' products and their pricing.

To visualize the extents of intra- and inter-substitution patterns among the manufacturers, we adapt and unfold the CV chart. The original operationalization of CV chart visualizes the category-level strength and weakness of each brand. That is, it focuses on and visualizes the focal brand's competitive clout and vulnerability with respect to the rest of the manufacturers as a whole, and fails to capture inter- and intra- manufacturer competitions. This may make sense in homogenous products market in which firms are typically assumed to produce single product (Kamakura and Russel 1984). However, manufacturers in differentiated products market are multi-product firms with a portfolio of diverse products. For them, it is important to understand the extents of intra- and inter-manufacturer substitution patterns. By unfolding the CV chart, we can visualize the extents of intra- and inter- manufacturer substitutions among the manufacturers. To that end, we follow and adjust the original definitions of CV chart. We define the clout of focal manufacturer $M$ with respect to another manufacturer $N$ as,

$$clout_{M,N} = \frac{1}{MN}\sum_{m=1}^{J_M}\sum_{n=1}^{J_N} e_{nm}^2, \tag{7}$$

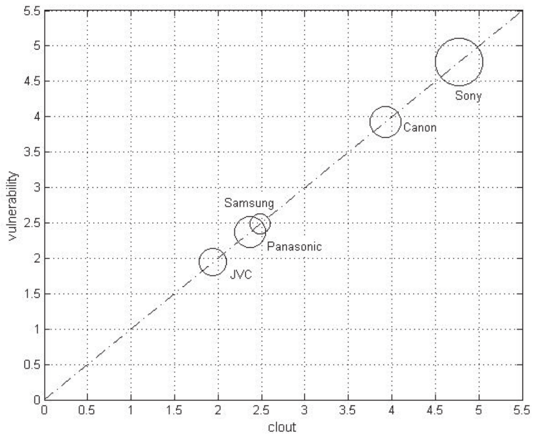
in which $m$ indexes products in $M$'s product line, $n$ does the same for $N$. $J_M$ is the number of products in $M$'s product line, $J_N$ the number of products in $N$'s product line, and $e_{nm}$ is the demand elasticity of $n$ with respect to $m$'s price change. Similarly, the vulnerability of focal manufacturer $M$ with respect to $N$ is defined as,

$$vulnerability_{M,N} = \frac{1}{MN}\sum_{n=1}^{J_N}\sum_{m=1}^{J_M} e_{mn}^2. \tag{8}$$

A short discussion is due. First, above equations express the average effect between a product in $M$ and another in $N$. Second, when $M \neq N$, above two equations express the extent of inter-manufacturer substitution. When $M = N$, the above two equations are identical and represent the extent of intra-manufacturer substitution. Lastly, note that $clout_{M,N} = vulnerability_{N,M}$.

The first step in developing the intra- and inter- CV charts is to compute the full cross price elasticity matrix. For this purpose, we use arc-elasticity formula in which we increase the price of the focal product by 10%, and simulate and compute the corresponding percentage market share changes for the rest of the products. Then, we use Equations of (7) and (8) and construct clout and vulnerability indices between a pair of manufacturers.

We show the intra-manufacturer CV chart in Figure 2. In this chart, X and Y values are all relative and their absolute levels do not matter. In the chart, each bubble represents a manufacturer and its size is proportional to its market share. Its X- and Y- values represent the average intra-manufacturer clout and vulnerability. A higher X (or Y) value means a higher level of substitution within the focal manufacturer among its customers. That is, if the focal manufacturer raises the price of one of its products, its consumers are more likely to substitute within the same manufacturer and less to other manufacturers. Alternatively, we may interpret that such manufacturers may have more loyal consumer base and hence more pricing power. From this Figure, we see that Sony has the highest levels of intra-brand clout and vulnerability, followed by Canon. JVC has the lowest level of intra-brand clout and vulnerability: JVC's customers are less likely to choose other JVC products if it raises its prices. Therefore, this chart is informative about relative pricing power among the manufacturers.

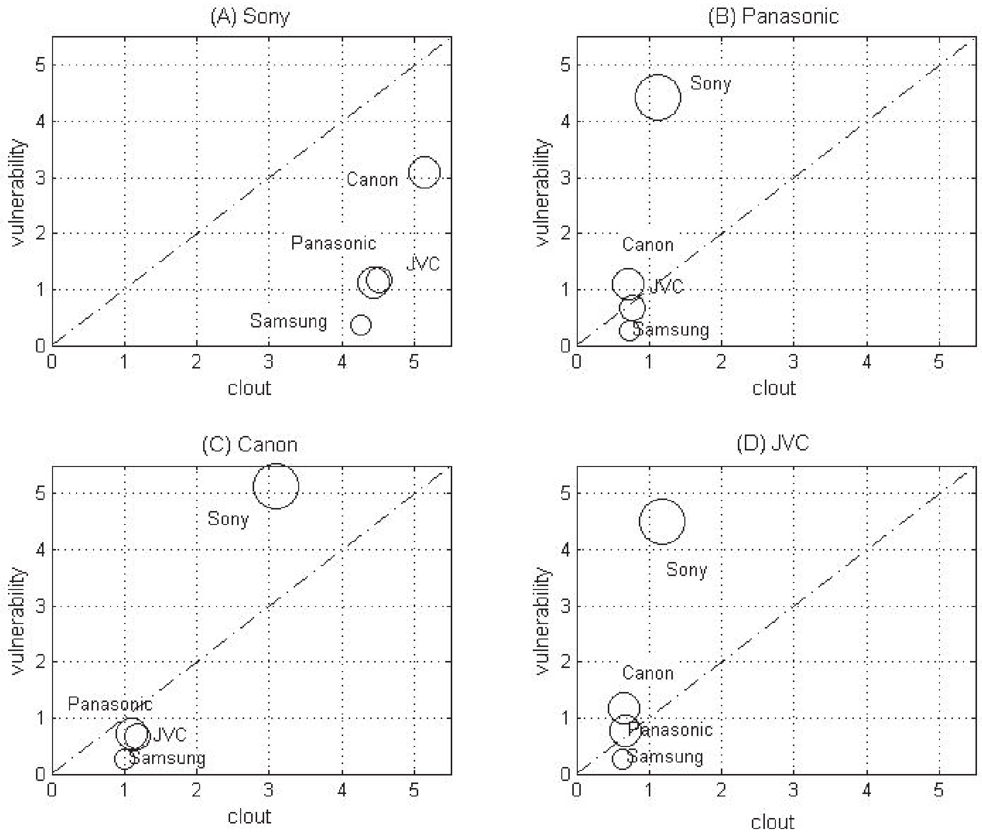



<Figure 2> Intra-manufacturer clout and vulnerability

We next discuss Figure 3.A and 3.B which show the inter- manufacturer clout and vulnerability between a pair of manufacturers. The focal manufacturer is found at the top of each panel. Panel (A) in Figure 3.A shows Sony's inter-manufacturer CV chart. In this panel, each bubble represents a manufacturer and its X and Y values are Sony's competitive clout and vulnerability with respect to other manufacturers. In general, we see that Sony's competitive clout levels are greater than its vulnerability levels against all other manufacturers since all the bubbles are placed below the 45-degree line. Among them, Canon is Sony's closest contender since Canon is positioned in the far upper right corner of Sony's inter-manufacturer CV chart. This means that Sony can gain the most from or lose the most to Canon by either party's price changes. However, Canon is still positioned below 45-degree line implying that Sony maintains an upper hand against Canon in the market. Note that Sony' clouts against all other manufacturers are all high while its vulnerability level is the lowest with Samsung. Therefore, we infer that Sony has the least to lose to Samsung in case Samsung lowers its price.

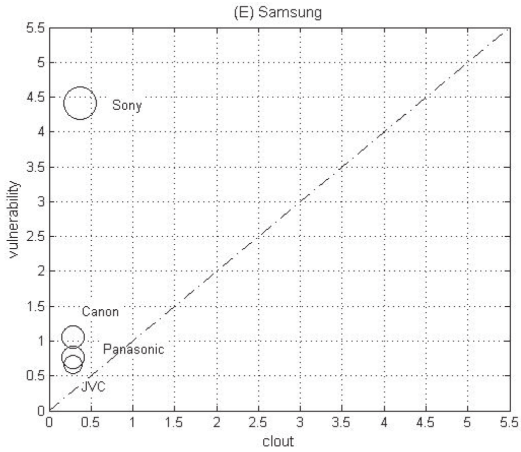
Inspecting the inter-manufacturer CV charts for other manufacturers as focal brands, we see that all other manufacturers have lowest clout and greatest vulnerability against Sony. From the panel (B) in Figure 3.A with Panasonic as the focal manufacturer, Sony is the most serious threat since Panasonic has the highest vulnerability to Sony. In contrast, Panasonic's competitive positions with the rest of the manufacturers are all similar although it is most competitive with respect to Samsung. From Figure 3.B, Samsung is the weakest manufacturer since all of the other manufacturers lie far above 45-degree line in its inter-manufacturer CV chart. In addition, Samsung may learn that it significantly lags behind Sony but not very much with the rest of the manufacturers in terms of product competition.

In summary, from the intra-manufacturer




〈Figure 3.A〉 Inter-manufacturer clout and vulnerability for selected manufacturers

CV charts, we conclude that Sony has the most pricing power among the manufacturers. From the inter-manufacturer CV charts, we conclude that Sony poses a uniform threat to all manufacturers in the competitive landscape and the rest of the manufacturers fiercely compete against one another. Compared to the typical operationalization of CV chart, we see that the intra- and inter-manufacturer CV charts unfold the conventional CV chart and offer a more detailed view on pairwise competitive positions among the multi-product manufacturers.




〈Figure 3.B〉 Inter-manufacturer clout and vulnerability for Samsung

## Ⅶ. Conclusion

In this paper, we study the intra- and inter-manufacturer competitions in differentiated products market using public data. To that end, we propose an approach to estimate consumer demand using sales rank data that are often publicly available from many online retail platforms. Leveraging the longitudinal nature of the data set, we adopt an alternative approach to account for the unobserved product characteristics in aggregate-level choice model. In the estimation framework, we first estimate the product specific intercepts in a multinomial logit and use them as additional product characteristics in random coefficient discrete choice model. This approach allows us to partly address the unobserved product characteristics in demand model.

As for the substantive application of the proposed approach, we unfold the clout and vulnerability chart and study consumers' intra- and inter-manufacturer substitution patterns among multi-product manufacturers in camcorder category. To that end, we simulate and estimate a very large scale full cross price elasticity matrix, which serves as an input to the unfolded CV chart. From the intra- and inter-manufacturer CV charts, we conclude that Sony shows the highest intra-substitution among consumers. This means that Sony's consumers switch within the brand in the presence of price hike and therefore Sony has the greatest pricing power among the manufacturers. From inter-manufacturer CV charts, we also find that Sony is the most serious threat to the rest of the manufacturers in our product category. In summary, the set of intra- and inter-manufacturer CV charts provides a more detailed understanding of the competition among multi-product manufacturers.

〈Received July 27. 2019〉
〈Accepted October 2. 2019〉

## References

Allenby, G. M. (1989). A unified approach to identifying, estimating and testing demand structures with aggregate scanner data. *Marketing Science*, 8(3), 265-280.

Bajari, P., Fox, J. T., & Ryan, S. P. (2008). Evaluating wireless carrier consolidation using semiparametric demand estimation. *Quantitative Marketing and Economics*, 6(4), 299.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association*, 90 (430), 443-450.

Bronnenberg, B. J., & Vanhonacker, W. R. (1996). Limited choice sets, local price response, and implied measures of price competition. *Journal of Marketing Research*, 33(2), 163-173.

Bruno, H. A., & Vilcassim, N. J. (2008). Research note—structural demand estimation with varying product availability. *Marketing Science*, 27(6), 1126-1131.

Chong, A. Y. L., Ch'ng, E., Liu, M. J., & Li, B. (2017). Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews. *International Journal of Production Research*, 55(17), 5142-5156.

DeSarbo, W. S., Grewal, R., & Wind, J. (2006). Who competes with whom? A demand-based perspective for identifying and representing asymmetric competition. *Strategic Management Journal*, 27(2), 101-129.

DeSarbo, W., & Rao, V. R. (1986). A constrained unfolding methodology for product positioning. *Marketing Science*, 5(1), 1-19.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904.

Guadagni, P. M., & Little, J. D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science*, 2(3), 203-238.

Hausman, J., Leonard, G., & Zona, J. D. (1994). Competitive analysis with differenciated products. *Annales d'Economie et de Statistique*, 159-180.

Howland, S (2018), Amazon beats Best Buy as top electronics retailer, *RetailDive.com*

Kadiyali, V., Sudhir, K., & Rao, V. R. (2001). Structural analysis of competitive behavior: New empirical industrial organization methods in marketing. *International Journal of Research in Marketing*, 18(1-2), 161-186.

Kamakura, W. A., & Russell, G. J. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26(4), 379-390.

Katahira, H. (1990). Perceptual mapping using ordered logit analysis. *Marketing Science*, 9(1), 1-17.

Kim, J. (2019). Aggregate Impact of Consumer Reviews on Market Outcome. *Working Paper.*

Knittel, C. R., & Metaxoglou, K. (2014). Estimation of random-coefficient demand models: two empiricists' perspective. *Review of Economics and Statistics*, 96(1), 34-59.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2), 132-157.

Lee, J., Kim, Y., Lee, J. D., & Park, Y. (2006). Estimating the extent of potential competition in the Korean mobile telecommunications market: Switching costs and number portability. *International Journal of Industrial Organization*, 24(1), 107-124.

MacKay, D. B., & Zinnes, J. L. (1986). A probabilistic model for the multidimensional scaling of proximity and preference data. *Marketing Science*, 5(4), 325-344.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2), 307-342.

Petrin, A., & Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, 47(1), 3-13.

Rutz, O. J., & Sonnier, G. P. (2011). The evolution of internal market structure. *Marketing Science*, 30(2), 274-289.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518-529.

Sudhir, K. (2001). Competitive pricing behavior in the auto market: A structural analysis. *Marketing Science*, *20*(1), 42-60.

Van Heerde, Harald J., Carl F. Mela, and Puneet Manchanda (2004). The dynamic effect of innovation on market structure. *Journal of Marketing Research*, 41(2), 166-183.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97-121.