JUN B. KIM, PAULO ALBUQUERQUE, and BART J. BRONNENBERG*

The authors propose a new method to visualize browsing behavior in so-called product search maps. Manufacturers can use these maps to understand how consumers search for competing products before choice, including how information acquisition and product search are organized along brands, product attributes, and price-related search strategies. The product search maps also inform manufacturers about the competitive structure in the industry and the contents of consumer consideration sets. The proposed method defines a product search network, consisting of the products and links that designate whether a product is searched conditional on searching other products. The authors model this network using a stochastic, hierarchical, and asymmetric multidimensional scaling framework and decompose the product locations as well as the product-level influences using product attributes. The advantages of the approach are twofold. First, the authors simultaneously visualize the positions of products and the direction of consumer search over products in a perceptual map of search proximity. Second, they explain the formation of the map using observed product attributes. The authors empirically apply their approach to consumer search of digital camcorders at Amazon.com and provide several managerial implications.

Keywords: brand networks, asymmetric multidimensional scaling, product search, hierarchical Bayes estimation

Mapping Online Consumer Search

The marketing literature has long recognized the importance of understanding the search phase of the consumer choice process for several reasons. First, prechoice consumer activities, such as search, reveal limits on consumer consideration sets (Roberts and Lattin 1991; Siddarth, Bucklin, and Morrison 1995; Urban, Hulland, and Weinberg 1993). Second, prepurchase product search reflects consumer strategies of information acquisition, including how search is organized, which may be informative about substitution patterns and potential choices. Third, knowledge of consumer information acquisition is fundamental to planning marketing communications and retail distribution (Newman and Staelin 1972). Finally, the introduction and widespread adoption of the Internet has greatly facilitated consumer information acquisition, and online consumer search has become ubiquitous. In this context, the goals of this article are to propose a practical and exploratory method that manufacturers can use to analyze and visualize rich consumer search patterns and obtain insights into the competitive structure of online markets in their industry.

Our modeling approach begins by defining a network representation of product search data. This network expresses the topology of search across products, such as whether a given product is searched conditional on another product being searched. Then, we analyze this network using asymmetric multidimensional scaling (MDS). In particular, we propose a hierarchical MDS model that estimates product positions in a latent attribute space and direction of search along pairs of products. In addition, we implement a property fitting regression (DeSarbo and Hoffman 1987) as a hierarchical regression step in a Bayesian estimation framework to interpret the dimensions of our latent attribute space. Our complete approach to describing the product search data yields a visualization that we refer to as a "prod-

*Jun B. Kim is Assistant Professor of Marketing, College of Management, Georgia Institute of Technology (e-mail: jun.kim@mgt.gatech.edu). Paulo Albuquerque is Assistant Professor of Marketing, Simon Graduate School of Business, University of Rochester (e-mail: paulo.albuquerque@simon.rochester.edu). Bart J. Bronnenberg is Professor of Marketing and CentER Fellow at CentER, Tilburg University (e-mail: bart.bronnenberg@uvt.nl). The authors thank Andrew Ainslie, Lee Cooper, Minha Hwang, Raphael Thomadsen, and Michel Wedel for discussion and advice during this project. The third author gratefully acknowledges support from the European Commission under Marie Curie Grant, IRG 230962. Michel Wedel served as associate editor for this article.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

uct search map." In this map, products likely to be searched together are located close to each other, and products unlikely to be searched together are placed at distant positions. The map also depicts the relative search attractiveness of each product and identifies the direction and asymmetry of consumer search among the products. Finally, the map can be used to shed some light on substitution patterns. Local subsets of products on the map can be interpreted as stereotypical products or consideration sets that are searched together and, presumably, compete more intensely. We also argue that the product search maps are an efficient and practical way to organize the massive number of possible consideration sets.

From a managerial perspective, this article offers three important and practical features to consumer durable goods manufacturers.[1] First, it provides a descriptive model of how products are searched online. From this model, manufacturers can obtain a detailed product-centric visualization of the competitive structure in their industry, revealing typical search patterns that involve their products. This visualization also helps them identify the set of competing products that are most frequently searched alongside their own products in the same session. Second, our proposed model has broad applicability. We apply our model to online search data from the largest online retailer in the world, Amazon.com. Although we illustrate the model with data from one product category, the method applies to many product categories sold at Amazon.com—notably, to durable goods sectors, for which useful data on search and sales are often difficult to find. Our method can also be applied to product search or site navigation data from several other large online retailers (e.g., Walmart.com). Third, although obtaining frequent product-level search or sales data using surveys is prohibitively expensive for firms selling durable goods in multiple categories, each involving a large number of alternatives, a benefit of our method is that it only uses publicly available summaries of search data. This makes our proposed descriptive method for studying consumer search cost effective. In addition to being free of charge, the data we use are not survey based but rather revealed measures of consumer search, which are viewed as more reliable (Newman and Lockeman 1975).

We empirically apply our method to the study of consumer search for digital camcorders at Amazon.com. From the analysis of our product search maps, we find that consumers predominantly organize their search for a camcorder by media format (e.g., DVD, hard disk, MiniDV); that is, consumers are more likely to search multiple products that share the same media format and less likely to search across media formats. Within each media format, consumer search is price driven, with similarly priced products more likely to be searched together and, thus, to be perceived as closer substitutes. Surprisingly, the brand attribute plays a less critical role than price or media format during the consumer search process in the camcorder category.

Finally, we demonstrate that manufacturers and product managers can use our estimated product search maps to conduct an in-depth, product-level analysis of consumer search. For each focal product, our results indicate the comprehensive set of comparison candidates, as well as the intensity of customer traffic to and from each of these candidates.

We organize the rest of the article as follows: In the next section, we discuss the relevant literature and subsequently describe the data used in this study. Then, we put forth our model, followed by a section on estimation. Next, we discuss the results of a numerical data experiment that verifies parameter recovery of the proposed stochastic MDS model. Following this, we discuss the results of the proposed model. Finally, we conclude with managerial implications and directions for further research.

RELEVANT LITERATURE

We discuss the two research streams most relevant to this study: (1) consideration set formation and information processing and (2) asymmetric MDS. In marketing, the consideration set literature follows the footsteps of economic theory of information search (Stigler 1961) because the concept of consideration is a logical outcome of information search (Hauser and Wernerfelt 1990; Roberts and Lattin 1991). In many previous two-stage choice models, researchers have inferred consideration sets from individual-level choice data (Bronnenberg and Vanhonacker 1996; Mehta, Rajiv, and Srinivasan 2003; Siddarth, Bucklin, and Morrison 1995; Swait and Erdem 2007), with the exception of Moe (2006), who uses individual-level clickstream data for browsed products. Common to these studies is that consideration sets are latent constructs inferred from the individual panel data in consumer packaged goods and that they improve the statistical fit of the empirical model (Chiang, Chibb, and Narasimhan 1998). This approach is not applicable to studying consumer durable goods, because repeat purchases are often too infrequently observed, making the empirical investigation of product substitution in durable goods challenging. We aim to overcome this challenge by exploiting information contents in product search by a large number of consumers.

In behavioral research, Payne (1976) and Bettman and Park (1980) report that the formation of consideration sets is associated with a subset of product attributes. Furthermore, Shugan (1980) shows that selective search can be a result of rational strategies of consumer search. Most relevant to a substantive contribution of the current study is Gilbride and Allenby's (2004) work, which investigates consumer use of attribute-based screening rules in a choice-based conjoint study. They infer the screening attributes and their importance from choice decisions and report that price and body style are the two most frequently used screening attributes in the camera category. Therefore, these are the attributes that camera manufacturers should primarily consider for new product development and planning marketing communications. However, it is reasonable to assume that screening attributes are category dependent. In this study, we propose a simple and cost-effective way to identify such patterns that emerge during consumer search in durable goods, using publicly available data for a large number of categories.

We also aim to add to the literature on brand mapping (e.g., Bijmolt and Wedel 1999; Elrod 1988; Erdem 1996).

[1]We emphasize that our target audience is not online retailers, which have access to more detailed data, but the manufacturers that would not have access to such data.

One aspect of our contribution is to visualize products in a latent attribute space, according to whether consumers tend to search for them together. This means that instead of estimating brand maps from choices or similarity ratings used in the past, we estimate such maps from search patterns. Moreover, our data on online search patterns are generally directional (i.e., search of product A given search of product B does not generally occur with equal incidence as in the opposite direction). For our modeling purpose, we adopt an MDS framework. Multidimensional scaling is a set of mathematical techniques that are suitable to uncover the latent structure among objects in a network by exploring similarities and dissimilarities in the data. Asymmetric MDS allows the similarity between two objects to be direction dependent. A few asymmetric MDS studies have previously appeared in the marketing literature. DeSarbo and Manrai (1992) operationalize Krumhansl's (1978) conceptual model of distance density and construct a visualization of the competitive automobile segments using switching data. DeSarbo, Grewal, and Wind (2006) propose a stochastic MDS model and analyze the asymmetric competitive market structure in luxury automobile and portable phone markets using consumer consideration and choice sets. Asymmetric MDS models have also been researched and used outside marketing (e.g., Okada and Imaizumi 1987; Saburi and Chino 2008). In these models, the symmetric part of the data is typically modeled with measures of interobject distance, such as Euclidean distance between product positions, and asymmetry is allowed by increasing distance in one direction but not the other. In the modeling section, we introduce a representation of the symmetric and asymmetric components adapted from Okada and Imaizumi (1987).

SEARCH DATA

Amazon.com summarizes and posts information from consumer prepurchase browsing activities in most durable goods categories. For each available product, the data show a list of products in the same product category that were frequently viewed by shoppers in one browsing session. For example, if a large number of consumers who viewed product j also viewed product k, k appears on the viewed product list (hereinafter view list) of j. Furthermore, products in j's view list are presented in descending order of frequency, with products that appear higher having a stronger relationship to product j than products appearing further down in the view list. Collected across all J products, we refer to these data as "product search data."

The product search data are an outcome of an item-to-item, collaborative filtering mechanism, in which the relationship between two products is determined by how frequently users jointly view the products (Linden, Smith, and Zada 2005). Appendix A, Part 1, presents details on Amazon.com's data generation. The product search data constitute a collection of directional relationships or links that exist between products. When considered together, these links lead to an associative network of choice alternatives, in which a node represents a product and an edge a relationship between two products. We represent the product search data using a $J \times J$ product search matrix Y, in which an entry y_{jk} represents a presence or absence of a relationship from product j to product k (i.e., $y_{jk} = 1$ if k appears on j's view list and $y_{jk} = 0$ if otherwise).[2] As Appendix A, Part 1, indicates, Amazon.com provides the top M most relevant products for a focal product in the view list. This means that $y_{jk} = 1$ if k is one of the top M most relevant products for focal j and $y_{jk} = 0$ if otherwise. These data are asymmetric because product j may be among the most relevant products for k, but not necessarily vice versa. Appendix A, Part 2, provides an illustrative example of the source of asymmetry in the product search data. The observed asymmetry in the product search data provides information to managers on how consumers navigate over choice options, thus allowing for a more comprehensive analysis of the relationships among the products. Therefore, it is our goal to incorporate such asymmetry explicitly in the proposed model.

For the empirical analysis, we used data from the digital camcorder category. Currently the dominant type in the camcorder category, digital camcorders store images and audio on a digital storage medium and offer good picture and sound qualities. In brief, our data collection process is as follows: We first downloaded Web pages for more than 250 camcorder products, each containing product-related data. For each product, two Web pages were downloaded. The first Web page contained information about product search data (i.e., which products consumers searched in the same online session as the chosen product). The second Web page contained detailed product information (e.g., list price, brand, media format, number of pixels, screen size, sales rankings, customer reviews), which we denote by product characteristics data. After we downloaded the Web pages, we parsed relevant information and assembled it into daily data sets. We repeated this process on a regular basis for a year beginning in June 2006 and constructed a longitudinal database.

Although our approach is scalable to the full set of products, for practical illustration, we narrowed down the number of products in the empirical study using the following criteria. First, we used products from the top four manufacturers (Sony, Canon, JVC, and Panasonic) and the three most common media formats (DVD, hard drive, and MiniDV). These four brands and three formats encompassed the large majority of digital camcorders available at Amazon.com during our data collection period. Second, we excluded professional grade digital camcorders because industry reports classify them as a separate category. Last, we excluded any product that does not appear in the view list of other products because consumer search of such product is not identified. Applying these criteria narrowed the number of products down to 62. All the top selling products are included in this subset.

In our empirical analysis, we used product search and characteristics data for August 2006. For these data, we define product j as being related to product k if j appears at least once on the daily view list of k during the month. We checked robustness of our analysis against alternative definitions of product relations. Specifically, we replicated our

[2]We do not have customer count data (i.e., the number of times two products are viewed together). Availability of such data would allow for a more detailed modeling approach, such as that of Wedel and DeSarbo (1996).

analysis with weighted relations between pairs of products, where the weights were based on the strength of product links over time. We operationalized this by converting the daily ranking into daily percentile rankings and then averaging them for August.[3] Our results are not sensitive to this alternative definition, and we do not include the details of this robustness check here to avoid repetition.

Table 1 shows the breakdown of these products by brand and storage media format. The table indicates that Sony has the most models sold at Amazon.com and that MiniDV is by far the most popular media storage format, constituting about half the products in our sample.

Table 2 presents a percentage breakdown of the relationships among the different brands. The first row shows the links of Sony products to products of other brands. For example, note that among all the relationships in the product search data, 18.51% are from Sony products to other Sony products and 5.17% are from Sony products to Panasonic products. For the reverse relationship, 6.25% of the links are from Panasonic to Sony.

We now discuss the relational information among the products in the product search data. Among the 62 products, the total number of existing relationships is 832, or 22% of all entries in the product search matrix Y, excluding the diagonal elements. The number of other products that appear on the view list of a given product ranges from 6 to 20, with a mean of 13.4 and a standard deviation of 3.2. The frequency with which a product appears on the view lists of other products ranges from 1 to 31, with a standard deviation of 7.5. The disparity between these two sets of summary statistics indicates that there are products that have short view lists but appear frequently on the view lists of other products (and vice versa), which provides evidence of asymmetry in the data.

Finally, we discuss a potential concern with the data available at Amazon.com. It is possible that Amazon.com could use the product search data to achieve its business goals (e.g., direct consumers to higher-margin products or clearance items). However, this is unlikely for several reasons. First, provision of truly similar products is strongly aligned with Amazon.com's commercial interests. By offering more relevant selections at lower search cost, Amazon.com is helping consumers choose products that best fit their needs, enhancing consumer shopping experience, and reducing price sensitivity (Lynch and Ariely 2000). Amazon.com's heavy investment in personalization and recommendation technologies reflects such interests. Second, the product search data are stable over time and do not show a sudden inclusion or radical movement of products at the top of the list, which would be expected if Amazon.com were to manipulate this list. Last, we verified through communications with a knowledgeable former manager at Amazon.com that the product summary data solely represent consumer browsing behavior.[4]

[3]We first encode the ranking of product r_{jt}, such that the most popular product at time t is encoded as J_t and not as 1. We then compute the percentile ranking of product j at time t as (Bajari, Fox, and Ryan 2008) $\hat{r}_{jt} = [r_{jt}/\max_k(r_{kt})] = (r_{jt}/J_t)$, where r_{jt} is j's ranking at t and J_t is the number of products at time t.

Table 1
NUMBER OF PRODUCTS BROKEN DOWN BY BRANDS AND MEDIA FORMATS

Formats	*Sony*	*Panasonic*	*Canon*	*JVC*	*Total*
MiniDV	8	8	11	6	33
Hard drive	4	0	0	7	11
DVD	9	5	4	0	18
Total	21	13	15	13	62

Table 2
PERCENTAGE BREAKDOWN OF PRODUCT SEARCH DATA AMONG BRANDS

Formats	*Sony*	*Panasonic*	*Canon*	*JVC*	*Total (%)*
Sony	18.51	5.17	6.13	3.85	33.66
Panasonic	6.25	8.53	5.05	1.44	21.27
Canon	6.49	5.53	13.34	1.80	27.16
JVC	4.45	2.88	3.00	7.57	17.90
Total (%)	35.70	22.11	27.52	14.66	100.00

Notes: For example, the cell (Panasonic, Sony) with a value of 6.25 indicates that 6.25% of all the relationships among the products are from Panasonic products to Sony products.

MODEL

The Asymmetric Distance Model

We chose to use an MDS approach, mainly driven by our interest in visualizing the similarity between product options that are searched online. Given the nature of our data, for which the tendency to search choice option k from option j is not the same as the tendency in the opposite direction, it is necessary to model similarity between a product pair asymmetrically. Therefore, our starting point in modeling is asymmetric MDS (DeSarbo and Manrai 1992; Holman 1979; Krumhansl 1978; Saburi and Chino 2008). This form of MDS generalizes symmetric MDS by using additional object-specific quantities that represent the skewness in otherwise symmetric distance. Specifically, following Saburi and Chino (2008), we define the asymmetric distance from product j to product k by

$$g_{jk} = d_{jk} - r_j + r_k, \tag{1}$$

where d_{jk} is the distance between product j and k and r_j and r_k are quantities to be estimated, which allow for $g_{jk} \neq g_{kj}$. We define d_{jk} as the Euclidean distance between two products j and k located at coordinates z_j and z_k in a derived latent attribute space of P dimension, as follows:

$$d_{jk} = \sqrt{\sum_{p=1}^{P}(z_{jp} - z_{kp})^2} = \|z_j - z_k\|. \tag{2}$$

The products j = 1, …, J can be represented graphically in maps by plotting their coordinates z_j in the latent attribute space. In this graph, r_j and r_k in Equation 1 can be depicted as the radii of circles centered at product j's and k's positions, respectively. This "position-circle" model is a

[4]Amazon.com chooses consumer trust and long-term relationship over short-term gains. Senior Amazon.com executives have also made this point to the press in the past. For example, the shareholder letter for the first quarter of 2009, as well as an earlier letter in 1997 available at Amazon.com's Web page, emphasizes the importance of long-term relationship with customers.

parsimonious way of representing the asymmetric similarities between brands. Thus, we model the symmetry in the data with the distance term, d_{jk}, whereas the asymmetry is captured by the difference in the radii, r_j and r_k. With this formulation, a product with a small radius will be searched more frequently and draw more search from large-radius products than vice versa.[5] A simple example illustrates this concept: Assume that for two hypothetical products, $d_{jk} = 5$, $r_j = 1$, and $r_k = 3$, where product k has a larger radius than product j. Computing distances using Equation 1, we obtain $g_{jk} = 7$ and $g_{kj} = 3$; that is, the distance from j to k is larger than that from k to j. We used this distance to express a greater likelihood (modeled next) of observing a link from the larger-radius product k to the smaller-radius product j than vice versa. To the extent that more search and more frequent consideration lead to higher sales, it is likely that a product with a small radius has higher demand than a large-radius product. To facilitate the analysis and interpretation of the map, we enforce strict positivity on the distance (d_{jk}) and radii (r_j and r_k) in Equation 1. In contrast, the resulting combined term of g_{jk} is not subject to such a restriction because our interpretation is not based on this term.

An implicit assumption in our approach is that the data are mainly driven by consumer search activities, and thus the searched products are informative about consumer desire to view the product pages. However, it is possible that Amazon.com's product recommendations and other online navigational tools influence consumer search. Among the many online features, we focus on product recommendations because research suggests that they affect consumer behavior the most (Bodapati 2008; Garfinkel et al. 2006). To infer a product search map net of the effects of Amazon.com's recommendations, we need to explicitly account for their effects. Our proposed approach is to control for such effects by adding observable recommendation features to Equation 1:

$$g_{jk} = d_{jk} - r_j + r_k - \beta X_{jk}^{rec}, \tag{3}$$

where X_{jk}^{rec} contains the online recommendations that relate options j and k and the coefficient β measures how much they affect the effective distance from j to k. In the presence of the last term, the effective distance g_{jk} is associated with the latent product positions, the radii, and Amazon.com's recommendations. Therefore, the estimated product locations z_j and radii r_j can be interpreted as net of the influence of recommendations. In the "Empirical Analysis" section, we provide details on how we operationalize the recommendations X_{jk}^{rec}.

To complete our model, we added stochasticity to the asymmetric distance variable g_{jk}. Stochastic MDS assumes that the effective distance between objects is obtained by the perturbation of the distance g_{jk}, which includes unexplained factors that may influence the relationship between j and k (DeSarbo and Cho 1989; Jedidi and DeSarbo 1991; Saburi and Chino 2008):

$$s_{jk} = g_{jk} + e_{jk} = d_{jk} - r_j + r_k - \beta X_{jk}^{rec} + e_{jk}, \tag{4}$$

where e_{jk} is a stochastic disturbance term. A feature of stochastic MDS models is that they allow for statistical inference of model parameters—in this case, the z_{jp}, r_j, and β, for all j = 1, …, J and p = 1, …, P.

In addition to plotting the product positions and accounting for asymmetry in product search, we aimed to interpret the dimensions of the derived space. To this end, we used property fitting, in which product characteristics such as brand and price are regressed on the estimated product positions, identifying for each characteristic a vector of directions in the map that represents the best-fitting relation to map positions. Intuitively, the obtained fitted vector points in the direction in which products with a given characteristic are located on the map.

Similar to property fitting, we interpreted the dimensions of the derived space using a hierarchical model, for which, at the upper level of the hierarchy, we estimated product positions and radii and, at the lower level of the hierarchy, we regressed the product positions and radii on product attributes. Property fitting involves regressing product attributes on map positions, whereas the hierarchical model proposed herein regresses map positions on product attributes (see also DeSarbo and Hoffman 1987). Specifically, the hierarchical model's purpose is to explain product positions z_{jp} as a linear model of product characteristics X_j:

$$z_{jp} = X_j\beta_p + \varepsilon_{jp},\ \varepsilon_{jp} \sim N(0, \sigma_p^2), \text{ for } \forall j, p, \tag{5}$$

where z_{jp} is the location of product j in dimension p, X_j is a [1 × K] vector of product j's attribute values, and β_p is a [K × 1] vector, where K is the number of product attributes. To be precise, Equation 5 represents the inverse of a property fitting regression. The regression coefficients $\beta_p = [\beta_{p,1}, \ldots, \beta_{p,K}]$ capture how the product attributes explain the product positions in the derived space. More specifically, β_p measures how well the product attributes X_j explain the pth dimension of the derived space. In a similar manner, we also explain product radii as a function of product attributes:

$$r_j = X_j\beta_r + \varepsilon_j,\ \varepsilon_j \sim N(0, \sigma_r^2), \text{ for } \forall j, \tag{6}$$

where β_r is a [K × 1] vector that measures the effects of product attributes on the magnitude of the radii.[6]

Likelihood Function

In our modeling strategy, we aim to explain the search patterns as represented by the product search matrix, Y, an

[5]There are alternative ways to implement the asymmetry. For example, DeSarbo and Manrai (1992) use a distance density model, in which r_j is modeled as the density of products around j. We allow the r_j to vary independently of density. We chose this proposed model for two reasons: First, the formulation is parsimonious and flexible enough to fully capture the asymmetry among the products, and second, we are interested in explaining the radii using observable product attributes.

[6]Instead of modeling the remaining product characteristics as explicitly affecting g_{jk} in Equation 4, we chose a hierarchical model in which the product characteristics implicitly affect g_{jk} through z_j and r_j. If we include the product characteristics in Equation 4 explicitly, the resulting map would only capture the residual relationship after accounting for the product characteristics. Therefore, the map would only reflect residual search behavior among the products. We believe that the proposed product search map, which captures all relevant product information, is much more managerially useful (as in DeSarbo and Jedidi 1995; DeSarbo and Rao 1986). In addition, there are two major differences between the proposed model and that of Hoff, Raftery, and Handcock (2002). First, our model explicitly addresses the asymmetry in the data. Second, we explain the latent positions and radii using the underlying product characteristics. Thus, our hyperparameters can be potentially used in policy simulations such as positioning and repositioning (DeSarbo and Hoffman 1987; DeSarbo and Jedidi 1995). This would be a challenge using Hoff, Raftery, and Handcock's (2002) approach.

unweighted and asymmetric matrix. An entry $y_{jk} = 1$ indicates a presence of a relationship from j to k, whereas $y_{jk} = 0$ indicates an absence of a relationship. The probability of observing a relationship from j to k using the effective distance from Equation 4 is expressed as follows:

$$(7) \quad \Pr(y_{jk} = 1) = \Pr(d_{jk} - r_j + r_k - \beta X_{jk}^{rec} + e_{jk} < \beta_0 + e_{jk0}),$$

where $\beta_0 + e_{jk0}$ is a random threshold for a link to be realized. Assuming that the error terms e_{jk} and e_{jk0} are i.i.d. random variables with an extreme value distribution, we can quantify the probability of observing a link between j and k as follows:

$$(8) \quad L(y_{jk} = 1|X^{rec}, z_j, z_k, r_j, r_k, \beta_0, \beta)$$
$$= \Pr(e_{jk} - e_{jk0} < \beta_0 - \|z_j - z_k\| + r_j - r_k + \beta X_{jk}^{rec}$$
$$= [1 + \exp(\|z_j - z_k\| - r_j + r_k - \beta X_{jk}^{rec} - \beta_0)]^{-1}.$$

Assuming conditional independence among relationships y_{jk} (see Hoff, Raftery, and Handcock 2002), the likelihood of the hierarchical model is given as follows:

$$(9) \quad L(Y|Z, R, \theta, X, X_{rec}) = \prod_{\forall j}\prod_{\forall k \neq j} \Pr(y_{jk}|X_{jk}^{rec}, z_j, z_k, r_j, r_k, \beta_0, \beta)$$
$$\times \Pr(z_j, z_k|\beta_p, \sigma_p^2, X)$$
$$\times \Pr(r_j, r_k|\beta_r, \sigma_r^2, X),$$

where $X = \{X_j\}$, $X^{rec} = \{X_{jk}^{rec}\}$, $Z = \{z_j\}$, $R = \{r_j\}$, and θ = the set of parameters. The conditional independence assumption means that after interproduct distances and asymmetries are accounted for, the relationships y_{jk} are independent. This model parsimoniously handles complex dependencies among the y_{jk} such as transitivity (i.e., in general, elevated search activity between products j and k and between products k and m implies elevated search activity between j and m) and reciprocity (i.e., elevated search activity from j to k is statistically associated with elevated search activity from k to j).[7]

ESTIMATION

Markov Chain Monte Carlo Estimation

We used Bayesian estimation to obtain the posterior distributions of the parameters of the proposed hierarchical model. Specifically, we used Markov chain Monte Carlo (MCMC) methods to simultaneously estimate the product positions z_j and radii r_j, as well as parameters β_p, β_r, and β. In Equation 3, we need to set one $r_j = 0$ for identification purposes (for additional details, see Appendix A, Part 3). Other identification conditions with respect to product positions are well documented by Abe (1998), Elrod (1988), and Erdem (1996). To initialize the MCMC, we estimated the product positions z_j and radii r_j using maximum likelihood estimation. After estimating positions z_j and radii r_j, we regressed them on the product characteristics and obtained the hyperparameter estimates of β_p and β_r. By using the maximum likelihood estimation values as the starting values in the MCMC, we aimed to reduce the number of burn-in iterations in the chain (Hoff, Raftery, and Handcock 2002). We used the Metropolis–Hastings algorithm and tuned the variances of the jumping distributions that generated candidate draws. To do so, we dynamically adjusted the variances of the jumping distributions during the burn-in period to achieve an acceptance ratio of approximately 20%–25%. We performed visual inspections of the chain to verify convergence. We drew 32,000 samples from the joint posterior distribution, using 27,000 iterations for burn-in and the final 5000 iterations to compute the posterior means and standard errors of the model parameters. Appendix B gives the detailed sampling sequence of the MCMC method and the prior distributions.

Data Experiment

To verify that the proposed model is well recovered, we designed and conducted a numerical data experiment. Our data experiment is based on the parameter estimates from the actual empirical data. That is, we first estimated the hypermodel parameters from the full product search model using the actual empirical data, treated them as the true model parameters, and then used them to generate the data for the experiment. We believe the proposed approach is more realistic than that based on a randomly chosen set of true parameters.

We generated the data as follows: With values of β_p, we stochastically determined the positions z_{jp} in a two-dimensional space (p = 1, 2) as a linear combination of the seven product attribute values (K = 7) in Equation 5. We also stochastically generated the radii r_j using β_r in Equation 6.

Given the product positions, radii, and other model parameters, we computed the asymmetric distance among all product pairs using Equation 3 and the corresponding link probability $\{p_{jk}\}$ of a product pair using Equation 8. Next, we created a binary matrix Y by performing Bernoulli trials using the computed probabilities $\{p_{jk}\}$, where Y is a realization from the underlying link probabilities.

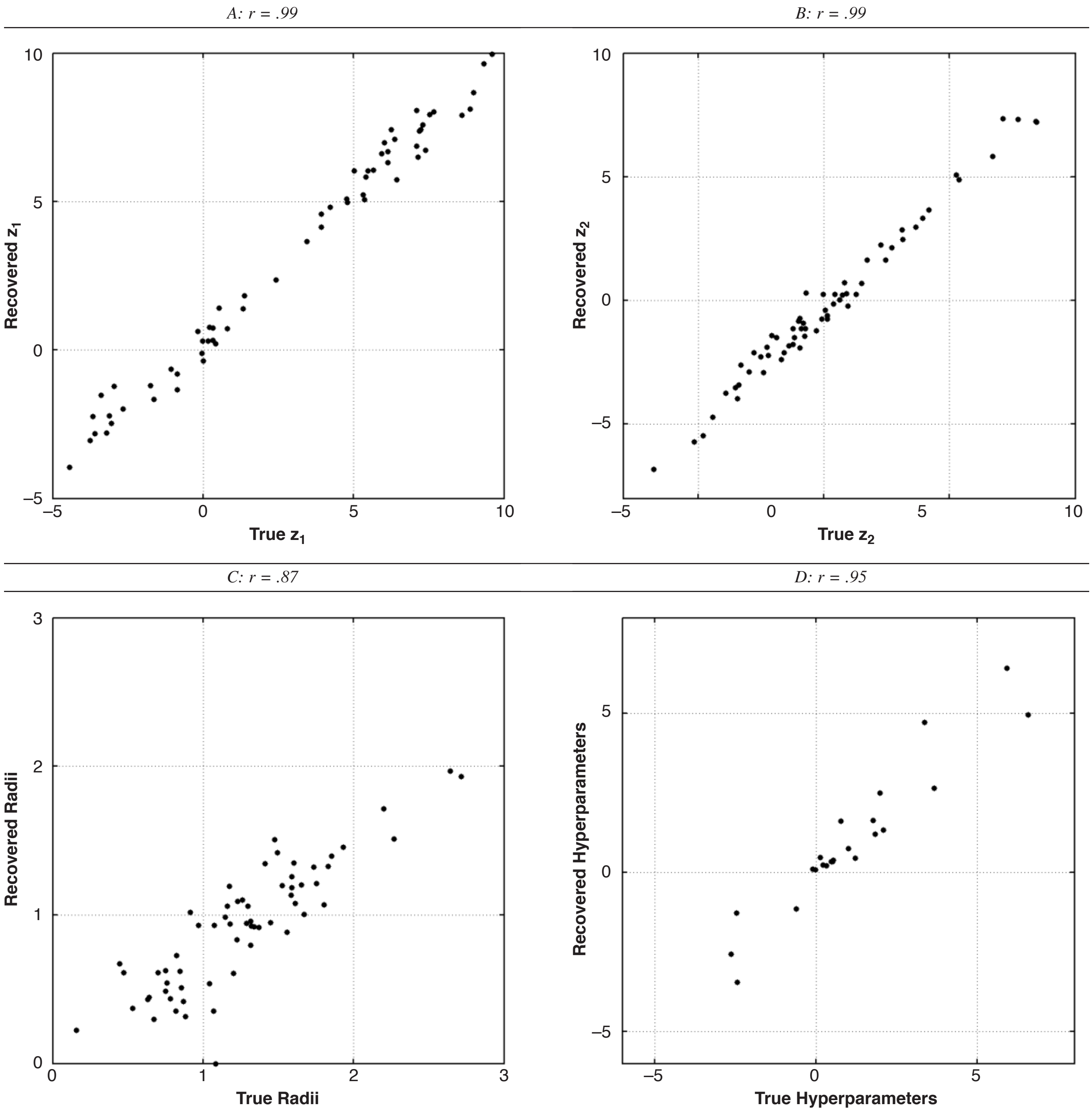
Next, using Y as our dependent variable, we estimated the hierarchical model using the steps outlined in the preceding section. The total number of parameters for location and radii is 186 because we estimate two coordinates and a radius for all J = 62 products. Only 182 parameters are identified because we needed to fix the location of one product, the radius of a product, and the first coordinate of another product. At every sweep of the sampler, in addition to sampling the locations and radii, we also drew from the posterior distributions of the hyperparameters.

The four subplots in Figure 1 facilitate a visual comparison of the true and recovered positions and radii for 62 products as well as the hyperparameters of the hierarchical model. We present the true parameter values on the x-axis and their recovered counterparts on the y-axis. In all subplots, we observed that the recovered parameters lie tightly centered on the 45-degree lines. Consistently, we find high correlations between the true and the mean of the recovered parameters, shown at the top of each subplot, ranging from .87 to .99. The 95% credibility interval of the posterior parameter distributions covers the majority of true parameter values. We conclude that the parameters of the proposed model are well recovered.

[7]Recall that in addition to modeling the links y_{jk}, we also modeled the strength of the link. We did this by modeling, for each product m, whether product j or product k appears higher on m's view list. This leads to a different model than the one in Equation 8 and a different likelihood function. Because the results from the two approaches are similar, we chose to use the parsimonious model described previously. The alternative formulation and results are available from on request.

Figure 1

DATA EXPERIMENT SCATTERPLOTS OF TRUE AND RECOVERED PRODUCT COORDINATES, RADII, AND HYPERPARAMETERS



Notes: Panels A, B, C, and D show first coordinates (z_1), second coordinates (z_2), radii (r), and other model parameters ($\{\beta_0, \beta, \beta_{rec}, \beta_{list}\}$), respectively. Correlations between the true and estimated parameters are shown at the top of each plot.

EMPIRICAL ANALYSIS

Model Selection

We used the deviance information criterion for model selection.[8] We estimated four models in a 2 × 2 selection design, varying the number of dimensions of the latent product space (two dimensions versus three dimensions) and the directionality of search (symmetric, with all radii equal to 0, versus asymmetric, with product-specific radii). The results are as follows:

	Two Dimensions	Three Dimensions
Symmetric	2005	1756
Asymmetric	1919	1598

[8]We also used Bayes factors (Kass and Raftery 1995) for model selection, computed from the log marginal density. This resulted in the same model selection.

For both the two- and three-dimensional models, the inclusion of radii improves the model fit significantly. Therefore, we conclude that the product search data warrant the explicit modeling of the asymmetry.

Comparing the specifications with two and three dimensions, we observe an improvement in the deviance information criterion with the additional dimension. Many MDS applications also report a similar statistical model fit improvement with the additional dimension to the derived space (DeSarbo and Manrai 1992). As previous literature has pointed out, the decision about which dimensionality to use for a given data set is as much substantive as statistical (Kruskal and Wish 1983). Following this convention, we chose the three-dimensional model over higher-dimensional models.

Effect of Online Recommendations

We now discuss how we operationalized Amazon.com's online recommendations X^{rec}_{jk} in Equation 3 and their influence on search. There are several forms of product recommendations at Amazon.com. Because our aggregate-level data summarize within-category consumer search activities, we are mainly concerned with recommendations for products within the same category. We identified two such features. First, Amazon.com provides recommendations to other same-category products based on purchases by past consumers conditional on viewing a product. These product recommendations provide easier access to relevant products, influence consumer search behavior, and may be responsible for the formation of interproduct relationships. Second, the default category page at Amazon.com provides a list of products sorted in terms of their popularity. Product proximity in this list may affect consumer behavior (Brynjolfsson, Dick, and Smith 2004), especially during the search initiation process, inducing consumers to conduct joint search among the options and creating interproduct relationships. We parameterized X^{rec}_{jk} in Equation 3 accordingly:

$$g_{jk} = d_{jk} - r_j + r_k - \beta X^{rec}_{jk} = d_{jk} - r_j + r_k - \beta_{rec}N_{jk} - \beta_{list}I_{jk}, \quad (10)$$

where N_{jk} is the fraction of days k is recommended from j, I_{jk} is an indicator variable that takes the value of 1 if j and k are located in the same page of the sorted product list, and the coefficients β_{rec} and β_{list} measure their respective effects on the effective distance g_{jk}. For example, $N_{jk} = .5$ means that k appears as a recommended product on j's product page 50% of the time during the data collection period. A larger N_{jk} and a positive β_{rec} imply a smaller g_{jk} because easy access from j to k may induce more consumer search of k from j. Similarly, if products j and k are located closer in the product list, $I_{jk} = 1$, they may lead to more frequent joint search of j and k, yielding a smaller distance g_{jk}.

In terms of results, we found that the posterior means of the recommendation coefficients, β_{rec} and β_{list}, are 10.96 and .38, with standard errors of 1.85 and .21, respectively. For products with recommendations (i.e., $\{j, k|N_{jk} > 0\}$ and $\{j, k|I_{jk} > 0\}$), these results imply that the Amazon.com recommendations increase link probabilities of $\Pr(y_{jk} = 1)$ by, respectively, .14 and .007 on average. We infer that both product recommendations and colocation in the sorted list have small but significant effects on the interproduct relationship. These findings are consistent with prior research (Garfinkel et al. 2006). The estimated latent product positions as well as the radii reported in the following section describe consumer search activities net of the effects of recommendations.[9]

Category-Level Consumer Product Search

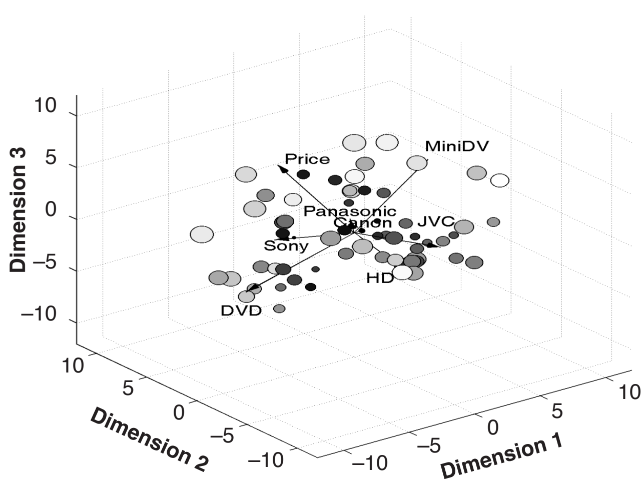
Figure 2 shows the posterior means of the product positions and radii. This figure depicts several pieces of information about search. First, consumers are more likely to search together products that are located closer to each other during online browsing sessions. The location of products in the derived space is not uniform, and Figure 2 suggests the existence of clusters of products. These clusters have managerial relevance to manufacturers, as we discuss subsequently, and our approach readily identifies cluster membership from the product search data.

Second, a small radius means that a product will be searched more frequently in relative terms because it is more likely to become a search destination from other products than a source for search to other products. Therefore, the radii depict the directions or flows of consumer search, which is from large- to small-radii products, and identify the "absorbing" products in consumer search (small-radii products). This implies that though consumers may initiate their searches in the area in which large-radii products are located, they tend to move toward and terminate search around the location of products with small radii. For better interpretation, we display sales ranks of the products in the figure as well, with darker circles indicating products with high sales. We find that the correlation between the sales ranks and the radii is .78, confirming that search and even-

[9] We also estimated Equation 10 without Amazon.com's recommendations. The correlation of pairwise effective distances (g_{jk}) between the models with and without the recommendations is .99. This high correlation implies that the recommendations have only marginal effect on the overall formation of the product search map because their occurrence is quite sparse.

Figure 2

POSTERIOR MEANS OF THE ESTIMATED PRODUCT LOCATIONS AND RADII IN THE PRODUCT SEARCH MAP



Notes: Circles represent product locations, circle radii represent the relative attractiveness of search, circle colors (darker) represent more sales, arrows represent the direction in which products with a given attribute can be found in the map, and the length of the arrows represents the degree of fit.

tually sales tend to be positively related to small-radii products. In addition, we observe that there are one or two products with small radii (and with higher sales) in each major cluster, which implies the existence of a small set of products that dominate search in the category (see also Figures 3 and 4).

Third, the arrows in the graph help interpret the latent attributes and represent the direction in which products with the corresponding attributes are likely to be found. For example, DVD-based products are more frequently found along the negative quadrant of the first dimension of the product search map. Technically, these arrows are constructed using property fitting regression coefficients. The arrow length is proportional to the R-square of each regression; in other words, a longer arrow means that the physical attribute is well explained by the latent dimensions. The arrow direction is computed as a ratio of the regression coefficients among the dimensions. Appendix A, Part 4, presents details on how to construct the arrows.

Consumer Search Patterns

We now analyze the resulting search patterns using Figures 3 and 4, the projections of the three-dimensional map onto two-dimensional spaces. First, we infer that media formats, which form three major segments in both maps, heavily guide consumer product search. This finding is confirmed by the property fitted arrows; the top three longest arrows, along with the price arrow in Figure 4, are all associated with the media formats. The three media format arrows are separated by angles of approximately 120 degrees in Figure 3, which is the maximum possible separation in a two-dimensional space. Their separations are also close to 120 degrees in Figure 4.

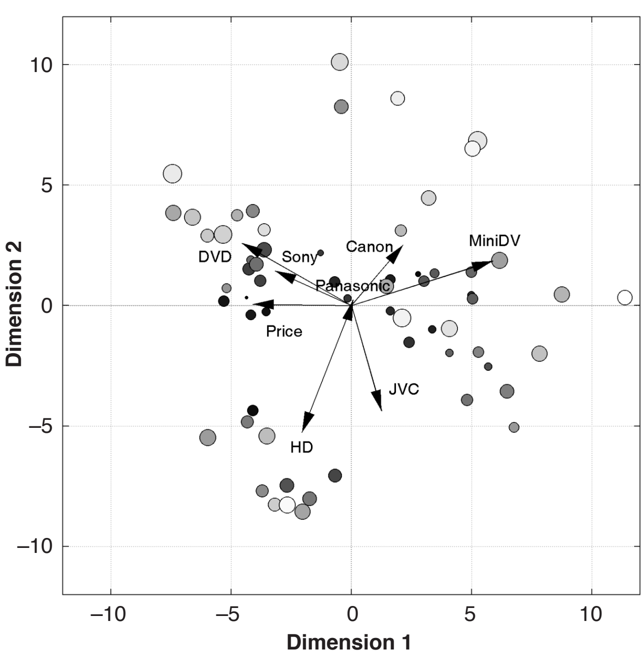
Second, from Figure 4, we note that the price explains the third (and also to some extent the first) dimension of the search map well. In this map, darker circles represent more expensive products (not higher sales, as in the previous two maps). The figure shows the transition of price from the lower-right-hand corner of the map (less expensive products) to the upper-left-hand corner (more expensive options). It strongly supports the view that the aggregate-level consumer search data contain information that identifies different segments in terms of price and that, in general, consumers search for products of similar price ranges.

Third, comparing the length of property fitted arrows for brand names with those of formats and price, we find that the effects of brand name on search are less important. More specifically, arrows for brand names such as Sony, JVC, and Canon are about half the size of the camcorder format arrows, and Panasonic's arrows are almost of zero length. Closer inspection reveals that Panasonic has products scattered all over the attribute space. In summary, the graph suggests that search takes place along media format first and, within media format, is based on price. In contrast, search is organized by brand to a lesser degree. We discuss the managerial implications of this search organization in the following sections.

Table 3 shows the values for the hyperparameters from our hierarchical regressions. They convey the same information as the regression coefficients from property fitting. In the hierarchical regression, one media format (DVD) and one brand (Panasonic) serve as the base, captured by the regression constant. Consistent with our findings from the

Figure 3

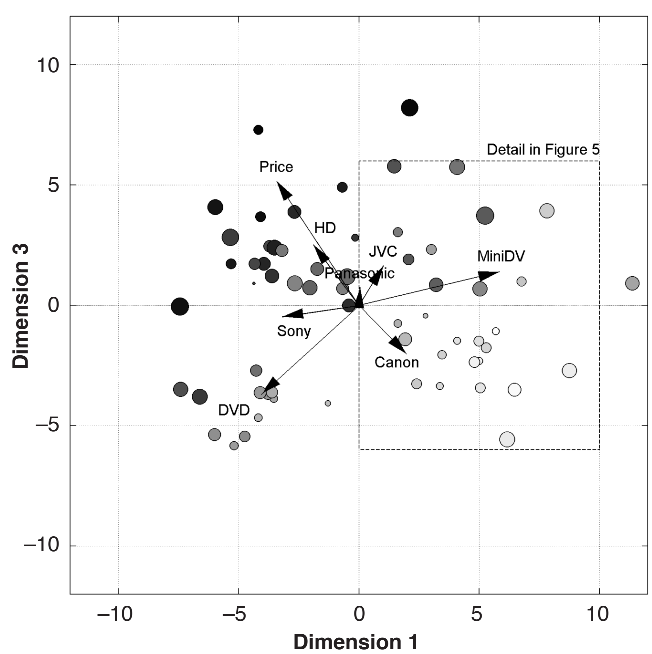
PROJECTION OF THREE-DIMENSIONAL PRODUCT SEARCH MAP ON DIMENSIONS 1 AND 2



Notes: The darker circles represent products of higher sales.

Figure 4

PROJECTION OF THREE-DIMENSIONAL PRODUCT SEARCH MAP ON DIMENSIONS 1 AND 3



Notes: The darker circles represent more expensive products.

Table 3

REGRESSION OF PRODUCT POSITIONS AND RADII ON PRODUCT ATTRIBUTES

	Dimension 1	*Dimension 2*	*Dimension 3*	*Radius*
Constant	3.54 (2.53)	–1.64 (–.92)	–9.43 (–8.73)	.49 (.85)
Sony	–.67 (–1.19)	1.70 (2.45)	–2.47 (–4.00)	–.08 (–.19)
Canon	1.18 (2.07)	2.71 (4.12)	–1.52 (–2.39)	.24 (.59)
JVC	2.21 (3.34)	–1.37 (–1.93)	–.57 (–.94)	.47 (.96)
MiniDV	6.95 (12.59)	–.50 (–.52)	3.86 (4.29)	–.01 (–.03)
Hard drive	.85 (.94)	–7.75 (–9.05)	3.17 (2.75)	–.07 (–.13)
Price	–6.97 (–4.07)	3.96 (1.98)	14.19 (10.48)	1.46 (1.71)
σ^2	2.88 (.56)	5.83 (1.12)	3.12 (.82)	.61 (.12)
R^2	.89	.71	.79	.20

Notes: t-statistics are in parentheses.

product search map, the most significant and important coefficients correspond to the two media formats and price.

Last, it is important to note that the inferred consumer search strategies are similar to those in recent empirical literature, even though we do not impose a priori restrictions on the nature of consumer search. When facing a large number of options, consumers adopt heuristics-based approaches, such as noncompensatory processes (Gilbride and Allenby 2004) and lexicographic strategies (Yee et al. 2007), using observed product attributes. Prior behavioral research has reported consumers' adoption of such rules (e.g., Bettman and Park 1980; Payne 1976). Our map shows that consumers use media formats and price to direct their search in the digital camcorder category.

Product-Level Competitive Analysis

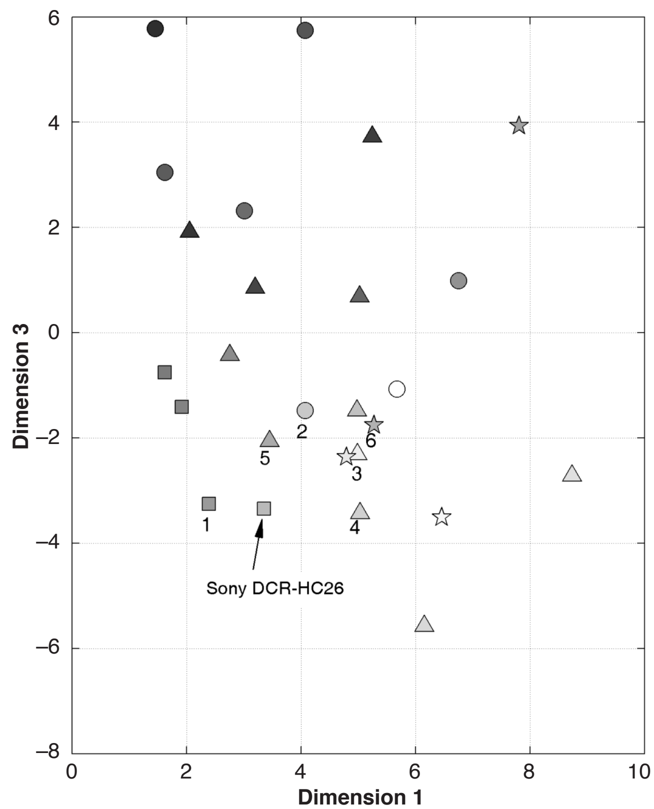
In the preceding section, we focused our discussion on consumer search patterns and competitive product structure at the category level. In this section, we demonstrate how manufacturers can use the product search map to identify neighboring products of each alternative in great detail.

As an illustration, Figure 5 focuses on an area with products in the MiniDV segment, which is located on the right-hand side of Figure 4. The products have different brands and show different prices. Darker colors indicate more expensive products, and brand names are coded with different symbols. Consistent with the conclusions of the previous section, we observe a clear pattern of increasing prices within the MiniDV segment. The majority of products at the bottom subcluster are priced lower than $300 and include the cheaper options in the MiniDV segment. Moving toward the top left, products become more expensive. Along this direction, different brands are scattered without a clear pattern. Therefore, we infer that within the MiniDV products, consumer search is more price driven, with similarly priced products more likely to be searched together and, thus, to be perceived as close substitutes. We note a similar pattern in the lower-left-hand corner for the DVD-based product cluster as well.

As a more specific example of how this map can be used, we now take the view of a manufacturer studying consumer search and the competitive structure in the category. Suppose that Sony is interested in establishing which products one of its models is regularly searched with and, thus, the products with which it closely competes. We focus on the Sony DCR-HC26, a MiniDV Digital Handycam Camcorder with 20× Optical Zoom priced at $307. Using Figure 5,

Figure 5

DETAIL OF THE PRODUCT SEARCH MAP IN THE MINIDV AREA



Notes: Plot symbols are as follows: Sony (square), Panasonic (circle), Canon (triangle), and JVC (star); darker colors indicate higher retail prices; and the numbered products constitute the six closest products to the Sony DCR-HC26.

Sony can identify and monitor the products located closest to this product. Table 4 lists the closest six competitors, along with some of their characteristics. These products are all MiniDVs and in a narrow price range. The average price of the cluster of numbered brands (including the Sony DCR-HC26) is $300. These products are the ones that Sony should monitor most closely because they often occur in the same set of searched products as the Sony DCR-HC26. In contrast, the average price of the cluster in the top left of the graph, containing two other Sony products, is $522. From

Table 4
PRODUCTS MOST CLOSELY SEARCHED JOINTLY WITH THE SONY DCR-HC26 RETAILING AT $307

Product Number	*Brand*	*Price ($)*	*Media*
1	Sony	333	MiniDV
2	Panasonic	292	MiniDV
3	Canon	262	MiniDV
4	Canon	285	MiniDV
5	Canon	313	MiniDV
6	JVC	310	MiniDV

Notes: For correspondence between product numbers and product positions, see Figure 5.

Table 5
MOST FREQUENTLY SEARCHED ALTERNATIVES FOR EACH TOP-SELLING PRODUCT

Focal Product/ Ranking of Conditional Search	*Brand*	*Price ($)*	*Media*
Sony DCR-DVD405/Retail price: $601			
1	Sony	743	DVD
2	Panasonic	699	DVD
3	Canon	647	DVD
4	Sony	522	DVD
5	Sony	639	MiniDV
Canon Elura 100/Retail price: $361			
1	Canon	313	MiniDV
2	Sony	394	MiniDV
3	Canon	293	MiniDV
4	Canon	262	MiniDV
5	Panasonic	430	MiniDV
Sony DCR-SR100/Retail price: $900			
1	Sony	600	Hard drive
2	Sony	700	Hard drive
3	Sony	800	Hard drive
4	JVC	596	Hard drive
5	JVC	697	Hard drive

the distance between the expensive and cheaper product clusters, the Sony DCR-HC26 is not often searched together with the more expensive products, suggesting that cannibalization of higher-priced and potentially higher-margin products is not a concern in this case.

Extending the preceding example, Table 5 lists the search sets (i.e., the most commonly searched alternatives) for the three best-selling products, two Sony and one Canon. First, we note that the search is aligned with the same media formats, with each product being searched often with the same format products. Second, price seems to play an important role as well, especially for the Canon Elura-100, because prices of the five closest search substitutes fall in a tight range of –$99 to +$69. In contrast, we observe a wider price range for products that closely compete with Sony DCR-SR100. Therefore, we can conclude that cannibalization is a concern for the relatively expensive Sony DCR-SR100 because the top three most jointly searched alternatives are other less expensive Sony products. Although we are not able to measure the level of cannibalization using our approach, we believe that this information is essential for Sony managers regarding its product line management.

In general, we conclude that the maps in Figures 2–5 are useful tools for manufacturers to gauge search patterns at one of the largest online retailers, including search proximity, direction of search, and organization of search. Moreover, given any focal product, the maps are an efficient medium for manufacturers to review the competitive structure among the products and identify likely consideration sets in which their products compete. In addition, the maps are based purely on publicly available data and are relatively easy to compute.

MANAGERIAL IMPLICATIONS AND CONCLUSIONS

To the best of our knowledge, no study has previously analyzed the structure of consumer information acquisition in a product category with many choice options. In the camcorder category, the set of searched products is typically a small subset of all options, and therefore knowledge about the contents of this subset is essential to understanding consumer preferences. Online browsing behavior forms a natural environment to study product search among durable goods. Our study is also the first to use site navigation data at the world's largest online retailer, Amazon.com, that are publicly available across many product categories to investigate prepurchase search patterns.

We model product search data using a hierarchical, stochastic, and asymmetric MDS model. Using hierarchical Bayesian estimation, we jointly estimate product positions and product-specific influence (i.e., radius), as well as hyperparameters that measure the contributions of product attributes on the formation of the dimensions of the product search map.

From a substantive perspective, the analysis of product search data in this article provides the following findings: First, using product search maps, managers can monitor in detail each product's neighboring competitors during consumer search stages (illustrated in Figure 5). This map enables managers to scrutinize the local relationships among products and, thus, to better understand substitution patterns for their own products during consumer search activities. Unlike many brand maps in previous marketing literature, which visualize competition among a few brands, this intuitive and informative map provides detailed information on whether a product is likely to be searched more often and if it is likely to be the end of the search process.

Second, a category-level analysis helps managers understand which and to what extent product attributes influence the contents of consumer product search. We provide several category-level insights, including the competitive product structure from the product positions, direction of consumer search from the product radii, and the effects of product attributes on the formation of the search map. In addition, the arrows in the product search map, resulting from the property fitting approach, depict the degree and direction of the influence of brands, media formats, and prices on the formation of consumer product search. Useful to marketing managers in the digital camcorder category is the finding that the intensity of consumer information search within the same media format is far greater than across media formats. We find that within the same media format, consumer search is more price driven than brand driven. Overall, we find that the role of brand is less significant. This finding is consistent with a recent, large-scale industry survey that reports the wavering power of brands in consumer electronics categories (Stewart-Allen/GMI Brand Barometer). This seems to be a finding with substantial implications in, for example, advertising content decisions. *Financial Times* recently quoted a Sony executive saying, "We cannot just rely on the brand to sell the product" in an article that reports Sony's planned attempt to shift its advertising strategy (Harding 2009). Our analysis using public data from 2006 shows signs that lead to a similar conclusion in the consumer electronics market.

Third, manufacturers can use the product search map to diagnose the performance of their product lines. For example, JVC has a MiniDV-based product offered at the retail price of $424 located at (17, 0, 0) in Figure 2. Judging from its isolated position on the map, JVC managers should infer that it is not searched frequently with other products and does not compete very effectively. The relative isolation of this JVC product is confirmed by the observation that the average distance between this specific JVC product and all other MiniDV-based products is 9.5, whereas the average distance among MiniDV-based products is 7.7. In turn, this information might cause JVC management to review its product to judge the efficacy of its positioning and viability. Last, we point out that as more online data are becoming available to practitioners and the marketing research community, well-tailored MDS techniques may prove to be useful exploratory tools to analyze and enhance understanding of brand search and consideration.

This study also has limitations, which may serve as further research opportunities. First, the collected data reflect the prepurchase browsing behavior of consumers, and thus the findings only apply to prepurchase stages of shopping behavior. However, given that subsets of searched products constitute consideration sets, from which final choices are made, our findings can be used with caution to infer mechanisms responsible for choice set generation in the digital camcorder category. Second, the nature of our data and model does not allow us to model consumer heterogeneity fully. Combining the current data with other sources, such as search frequency or choice data, may make such an investigation possible. We leave this for further research.

APPENDIX A

Part 1: Amazon.com's Aggregate-Level Product Search Data

According to Amazon.com's patent (Linden, Smith, and Zada 2005), the sequence of operations generating the product search data is as follows:

1. User clickstream or query log data that reflect products each user views during an ordinary browsing session are stored for a certain period. A product is shown to a shopper only if the corresponding product detail page is requested.
2. The normalized degree of relationship between two products is measured according to how frequently consumers view them together using $rjk = (n_{jk}/\sqrt{n_j}\sqrt{n_k})$, where n_j is the number of consumers who searched product j and n_{jk} is the number of consumers who searched j and k.
3. The preceding measurement is repeated on all pairs of products.
4. For each focal product, related products are sorted in the order of a descending relationship.
5. Among the sorted products, products outside the focal product's category are removed from the list. Note that a category can be defined in several different ways at Amazon.com. We only collected search data in the camcorder industry for choice options for digital camcorders and analog camcorders based on the Hi8 medium.
6. The top M related products are extracted for each focal product.

Part 2: Illustration for Asymmetric Product Search Data

We illustrate the asymmetry in the product search data. Assume three products, A, B, and C, and the following numbers of consumers who viewed each product and pairs of products: $n_A = 20$, $n_B = 10$, $n_C = 10$, $n_{AB} = 5$, $n_{BC} = 3$, and $n_{AC} = 4$.

The relationships among the three products are computed as follows:

$$r_{AB} = \frac{n_{AB}}{\sqrt{n_A}\sqrt{n_B}} = .35, r_{BC} = \frac{n_{BC}}{\sqrt{n_B}\sqrt{n_C}} = .30, r_{AC} = \frac{n_{AC}}{\sqrt{n_A}\sqrt{n_C}} = .28.$$

Table A1 lists the related products for each product using the preceding computed quantities. The first column represents the focal products and the first row represents the related products. The number in parentheses is the order in which the products appear in the view list.

Now, we focus on the one product in each row that is most closely related to each focal product. Products A's, B's, and C's closest relationships are to products B, A, and B, respectively. Table A1 shows that the product pair (A, B) is symmetric because they have the closest relationship to each other. The pair (A, C) is also symmetric because B is the closest to both products. However, the pair (B, C) is asymmetric because A is closest to B but B is closest to C; they are not mutually closest. For product search data involving a large number of products, there will be symmetric as well as asymmetric pairs of products, reflecting complex consumer product search behaviors.

Table A1
ILLUSTRATION FOR ASYMMETRIC PRODUCT SEARCH DATA

Focal Products	*A*	*B*	*C*
A		.35 (1)	.28 (2)
B	.35 (1)		.30 (2)
C	.28 (2)	.30 (1)	

Part 3: Identification Restrictions

The identification restrictions for our three-dimensional model are as follows: We adopt the identification restrictions in MDS that Abe (1998), Elrod (1988), Erdem (1996), and Okada and Imaizumi (1987) outline. The main goal is to prevent translation, rotation, and refection of positions during the estimation process. The conditions are as follows:

1. One product located at (0, 0, 0).
2. One product located at (0, 0, R+).
3. One product located at (0, R+, R+).

Note that in a T-dimensional space, there are T degrees of freedom for translation, T(T – 1)/2 degrees of freedom for rotation, and T degrees of freedom for reflection. Therefore, for T = 3, we must apply nine constraints in total. In addition, to identify the radii, we set the radius of one product to a constant (Okada and Imaizumi 1987).

Part 4: Property Fitting

In this section, we discuss how we obtain the property fitting arrows in the product search maps in Figures 2, 3, and

4. After we estimate the product positions $\mathbf{Z}$, we regress each product attribute on the product positions,

$$X_k = \mathbf{Z}\beta_k + E,$$

where X_k is a [J × 1] vector for the kth attribute values (k = 1, …, K) for J products; $\mathbf{Z}$ is a [J × P] matrix, with each product j's coordinates $[z_{j1}, \ldots, z_{jP}]$ on row j; β_k is a [P × 1] vector that measures the contribution of $\mathbf{Z}$ to X_k; and E is a [J × 1] vector of errors. In our empirical analysis, J = 62, P = 3, and K = 7. For each regression for k, we obtain R_k^2 and $\hat{\beta} = [\beta_{k,1}, \beta_{k,2}, \beta_{k,3}]$, which jointly determine the arrow in Figure 2. If the kth attribute values are well explained by $\mathbf{Z}$, we obtain a high R_k^2. To represent this scenario with a long arrow, we compute the arrow vector for the kth attribute as follows:

$$\left(R_k^2 \frac{\beta_{k,1}}{\sqrt{\beta_{k,1}^2 + \beta_{k,2}^2 + \beta_{k,3}^2}}, R_k^2 \frac{\beta_{k,2}}{\sqrt{\beta_{k,1}^2 + \beta_{k,2}^2 + \beta_{k,3}^2}}, R_k^2 \frac{\beta_{k,3}}{\sqrt{\beta_{k,1}^2 + \beta_{k,2}^2 + \beta_{k,3}^2}} \right).$$

Note that the total length of this arrow is R_k^2 (as desired). We repeat this process for all k to construct the arrows shown in the map.

APPENDIX B

The following are the priors for the hypermodel parameters:

$$\beta_p \sim N\left(\beta_p^0, V_{\beta_p}^0\right), \sigma_p^2 \sim IG\left(\upsilon_p^0, s_p^0\right), V_z = \begin{bmatrix} \ddots & 0 & 0 \\ 0 & \sigma_p^2 & 0 \\ 0 & 0 & \ddots \end{bmatrix}, \text{and}$$

$$\beta_r \sim N\left(\beta_r^0, V_{\beta_r}^0\right), \sigma_r^2 \sim IG\left(\upsilon_r^0, s_r^0\right).$$

Furthermore, we assume the following parameters for the priors K = 7; p = 1, …, P; and P = 3:

$$\beta_p^0 = 0_K, V_{\beta_p}^0 = 10^6 I_K, \upsilon_p^0 = 4, s_p^0 = 1, \beta_r^0 = 0_K, V_{\beta_r}^0 = 10^6 I_K,$$

$$\upsilon_r^0 = 4, s_r^0 = 1.$$

Here, I_K represents an identity matrix of size [K × K], and 0_K is a vector [K × 1] of zeros. The following is the sequence for the Gibbs sampler:

Step 1. Draw β_p, p = 1, …, P.

$$\beta_p \middle| X, Z_p, \sigma_p^2, \beta_p^0, V_{\beta_p}^0$$

$$\sim N\left\{\left[\left(V_{\beta_p}^0\right)^{-1} + X'X\right]^{-1}\left[\left(V_{\beta_p}^0\right)^{-1}\beta_p^0 + X'Z_p\right], \left[\left(V_{\beta_p}^0\right)^{-1} + X'X\right]^{-1}\right\},$$

where $Z_p = [z_{1p}, \ldots, z_{Jp}]$ and X is a [J × K] matrix with each product j's attributes $X_j = [X_{j1}, \ldots, X_{jK}]$ on row j.

Step 2. Draw σ_p^2, p = 1, …, P.

$$\sigma_p^2 \middle| X, Z_p, \beta_p, \upsilon_p^0, s_p^0 \sim IG\left(\upsilon_p^0 + J, \frac{\upsilon_p^0 s_p^0 + J\bar{s}_p}{\upsilon_p^0 + J}\right),$$

where $Z_p = [z_{1p}, \ldots, z_{Jp}]'$ and $\bar{s}_p = (1/J)\Sigma_j(z_{jp} - X_j\beta_p)^2$.

Step 3. Draw β_r.

$$\beta_r \middle| X, R, \sigma_r^2, \beta_r^0, V_{\beta_r}^0$$

$$\sim N\left\{\left[\left(V_{\beta_r}^0\right)^{-1} + X'X\right]^{-1}\left[\left(V_{\beta_r}^0\right)^{-1}\beta_r^0 + X'R\right], \left[\left(V_{\beta_r}^0\right)^{-1} + X'X\right]^{-1}\right\},$$

where $R = [r_1, \ldots, r_J]$.

Step 4. Draw σ_r^2.

$$\sigma_r^2 \middle| X, R, \beta_r, \upsilon_r^0, s_r^0 \sim IG\left(\upsilon_r^0 + J, \frac{\upsilon_r^0 s_r^0 + J\bar{s}_r}{\upsilon_r^0 + J}\right),$$

where $\bar{s}_r = (1/J)\Sigma_j(r_j - X_j\beta_r)^2$.

Step 5. Draw z_j.

Product j's position is $z_j = [z_{j1}, \ldots, z_{jp}]$. The conditional posterior, K, for z_{jp} is

$$K\left(z_{jp} \middle| Y, N_{rec}, I_{list}, Z_{-jp}, \{r_j\}, X_j, \beta_p, \sigma_p^2, \beta_{rec}, \beta_{list}\right)$$

$$\propto L\left(Y \middle| N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right) \times p\left(z_{jp} \middle| X_j, \beta_p, \sigma_p^2\right),$$

where $Z_{-jp} = \{z_j\}\backslash z_{jp}$, $Y = \{y_{jk}\}$, $N_{rec} = \{N_{jk}\}$, $I_{list} = \{I_{jk}\}$, and

$$L\left(Y \middle| N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right)$$

$$= \prod_{j=1}^{J} \prod_{k \neq j}^{J} \Pr\left(y_{jk} = 1 \middle| N_{rec}, I_{list}, z_j, z_k, r_j, r_k, \beta_{rec}, \beta_{list}\right),$$

$$p\left(z_{jp} \middle| X_j, \beta_p, \sigma_p^2\right) = \phi\left(X_j\beta_p, \sigma_p^2\right),$$

where ϕ is the probability density function of the normal distribution. We draw from the posterior using a Metropolis–Hastings algorithm. We use a normal distribution as the proposal distribution, $p(z'_{jp}|z_{jp}) = \phi(z_{jp}, \sigma_z^2)$. We accept the candidate, z'_{jp}, with the following acceptance probability:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, Z_{-jp}, z'_{jp}, \{r_j\}, \beta_{rec}, \beta_{list}\right)}{L\left(Y|N_{rec}, I_{list}, Z_{-jp}, z_{jp}, \{r_j\}, \beta_{rec}, \beta_{list}\right)} \times \frac{p\left(z'_{jp}|X_j, \beta_p, \sigma_p^2\right)}{p\left(z_{jp}|X_j, \beta_p, \sigma_p^2\right)} \times \frac{p\left(z'_{jp}|z_{jp}\right)}{p\left(z_{jp}|z'_{jp}\right)}\right].$$

Because ϕ is symmetric, the ratio $p(z'_{jp}|z_{jp})/p(z_{jp}|z'_{jp})$ is unity. Therefore, we can simplify the preceding acceptance probability as follows:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, Z_{-jp}, z'_{jp}, \{r_j\}, \beta_{rec}, \beta_{list}\right)}{L\left(Y|N_{rec}, I_{list}, Z_{-jp}, z_{jp}, \{r_j\}, \beta_{rec}, \beta_{list}\right)} \times \frac{p\left(z'_{jp}|X_j, \beta_p, \sigma_p^2\right)}{p\left(z_{jp}|X_j, \beta_p, \sigma_p^2\right)}\right].$$

Step 6. Draw r_j.

The conditional posterior for r_j is as follows:

$$K\left(r_j|Y, N_{rec}, I_{list}, \{z_j\}, R_{-j}, X_j, \beta_r, \sigma_r^2, \beta_{rec}, \beta_{list}\right)$$
$$\propto L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right) \times p\left(r_j|X_j, \beta_r, \sigma_r^2\right),$$

where $R_{-j} = \{r_k\}\backslash r_j$ and $p(r_j|X_j, \beta_r, \sigma_r^2) = \phi(X_j\beta_r, \sigma_r^2)$.

We accept the candidate r'_j with the following acceptance probability:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, \left\{z_j\right\}, R_{-j}, r'_j, \beta_{rec}, \beta_{list}\right)}{L\left(Y|N_{rec}, I_{list}, \left\{z_j\right\}, R_{-j}, r_j, \beta_{rec}, \beta_{list}\right)} \times \frac{p\left(r'_j|X_j, \beta_r, \sigma_p^2\right)}{p\left(r_j|X_j, \beta_r, \sigma_p^2\right)} \times \frac{p\left(r'_j|r_j\right)}{p\left(r_j|r'_j\right)}\right],$$

where we use the normal distribution as the proposal distribution $p(r'_j|r_j) = \phi(r_j, \sigma_R^2)$. Because ϕ is symmetric, the ratio $p(r'_j|r_j)/p(r_j|r'_j)$ is unity. Therefore, we can further simplify the acceptance probability as follows:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, \left\{z_j\right\}, R_{-j}, r'_j, \beta_{rec}, \beta_{list}\right)}{L\left(Y|N_{rec}, I_{list}, \left\{z_j\right\}, R_{-j}, r_j, \beta_{rec}, \beta_{list}\right)} \times \frac{p\left(r'_j|X_j, \beta_r, \sigma_p^2\right)}{p\left(r_j|X_j, \beta_r, \sigma_p^2\right)}\right],$$

Step 7. Draw β_{rec}.

The conditional posterior for β_{rec} is as follows:

$$K\left(\beta_{rec}|Y, N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{list}\right)$$
$$\propto L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right) \times p\left(\beta_{rec}\right),$$

where $p(\beta_{rec})$ is the prior distribution for β_{rec}. We accept the candidate β'_{rec} with the following acceptance probability:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta'_{rec}, \beta_{list}\right)}{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right)} \times \frac{p\left(\beta'_{rec}\right)}{p\left(\beta_{rec}\right)} \times \frac{p\left(\beta'_{rec}|\beta_{rec}\right)}{p\left(\beta_{rec}|\beta'_{rec}\right)}\right].$$

Because we assume a diffuse prior for $p(\beta_{rec})$ and the normal probability density function for the proposal distribution of $p(\beta'_{rec}|\beta_{rec})$, both ratios of $p(\beta'_{rec})/p(\beta_{rec})$ and $p(\beta'_{rec}|\beta_{rec})/p(\beta_{rec}|\beta'_{rec})$ are unity. Therefore, we can simplify the preceding acceptance probability as follows:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta'_{rec}, \beta_{list}\right)}{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right)}\right].$$

Step 8. Draw β_{list}.

The conditional posterior for β_{list} is as follows:

$$K\left(\beta_{list}|Y, N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}\right)$$
$$\propto L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right) \times p\left(\beta_{list}\right),$$

where $p(\beta_{list})$ is the prior distribution for β_{list}. We accept the candidate β'_{list} with the following acceptance probability:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta'_{list}\right)}{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right)} \times \frac{p\left(\beta'_{list}\right)}{p\left(\beta_{list}\right)} \times \frac{p\left(\beta'_{list}|\beta_{list}\right)}{p\left(\beta_{list}|\beta'_{list}\right)}\right].$$

Because we assume a diffuse prior for $p(\beta_{list})$ and the normal probability density function for the proposal distribution of $p(\beta'_{list}|\beta_{list})$, both ratios of $p(\beta'_{list})/p(\beta_{list})$ and $p(\beta'_{list}|\beta_{list})/p(\beta_{list}|\beta'_{list})$ are unity. Therefore, we can simplify the preceding acceptance probability as follows:

$$\min\left[1, \frac{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta'_{list}\right)}{L\left(Y|N_{rec}, I_{list}, \{z_j\}, \{r_j\}, \beta_{rec}, \beta_{list}\right)}\right].$$

REFERENCES

Abe, Makoto (1998), "Error Structure and Identification Condition in Maximum Likelihood Nonmetric Multidimensional Scaling," *European Journal of Operational Research*, 111 (2), 216–27.

Bajari, Patrick, Jeremy T. Fox, and Stephen P. Ryan (2008), "Evaluating Wireless Carrier Consolidation Using Semiparametric Demand Estimation," *Quantitative Marketing and Economics*, 6 (4), 299–338.

Bettman, James R. and C. Whan Park (1980), "Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Analysis: A Protocol Analysis," *Journal of Consumer Behavior*, 7 (3), 234–48.

Bijmolt, Tammo and Michel Wedel (1999), "A Comparison of Multidimensional Scaling Methods for Perceptual Mapping," *Journal of Marketing Research*, 36 (May), 277–85.

Bodapati, Anand (2008), "Recommendation Systems with Purchase Data," *Journal of Marketing Research*, 45 (February), 77–93.

Bronnenberg, Bart J. and Wilfried R. Vanhonacker (1996), "Limited Choice Sets, Local Price Response, and Implied Measures of Price Competition," *Journal of Marketing Research*, 33 (May), 163–73.

Brynjolfsson, Erik, Astrid A. Dick, and Michael D. Smith (2004), "Search and Product Differentiation at an Internet Shopbot," working paper, Sloan School of Management, Massachusetts Institute of Technology.

Chiang, Jeongwen, Siddhartha Chibb, and Chakravarthi Narasimhan (1998), "Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity," *Journal of Econometrics*, 89 (1–2), 223–48.

DeSarbo, Wayne S. and Jaewun Cho (1989), "A Stochastic Multidimensional Scaling Vector Threshold Model for the Spatial Representation of 'Pick Any/n' Data," *Psychometrika*, 54 (1), 105–129.

———, Rajdeep Grewal, and Jerry Wind (2006), "Who Competes with Whom? A Demand-Based Perspective for Identifying and Representing Asymmetric Competition," *Strategic Management Journal*, 27 (2), 101–129.

——— and Donna L. Hoffman (1987), "Constructing MDS Joint Spaces from Binary Choice Data: A Multidimensional Unfolding Threshold Model for Marketing Research," *Journal of Marketing Research*, 24 (February), 40–54.

——— and K. Jedidi (1995), "The Spatial Representation of Heterogeneous Consideration Sets," *Marketing Science*, 14 (3), 326–42.

——— and Ajay K. Manrai (1992), "A New Multidimensional Scaling Methodology for the Analysis of Asymmetric Proximity Data in Marketing Research," *Marketing Science*, 11 (1), 1–20.

——— and V.R. Rao (1986), "A Constrained Unfolding Methodology for Product Positioning," *Marketing Science*, 5 (1), 1–19.

Elrod, Terry (1988), "Choice Map: Inferring a Product-Market Map from Panel Data," *Marketing Science*, 7 (1), 21–40.

Erdem, Tülin (1996), "A Dynamic Analysis of Market Structure Based on Panel Data," *Marketing Science*, 15 (4), 359–78.

Garfinkel, Robert, R.D. Gopal, Bhavik K. Pathak, Rajkumar Venkatesan, and Fang Yin (2006), "Empirical Analysis of the Business Value of Recommender Systems," working paper, School of Business, University of Connecticut.

Gilbride, Timothy J. and Greg M. Allenby (2004), "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," *Marketing Science*, 23 (3), 391–406.

Harding, Robin (2009), "Sony Retunes $5bn Ad Budget," *Financial Times*, (November 17), (accessed October 8, 2010), [available at http://www.ft.com/cms/s/2/0315c95c-d3a9-11de-8caf-00144feabdc0.html].

Hauser, John R. and Birger Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16 (4), 393–408.

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock (2002), "Latent Space Approaches to Social Network Analysis," *Journal of the American Statistical Association*, 97 (Winter), 1090–1098.

Holman, E.W. (1979), "Monotonic Models for Asymmetric Proximities," *Journal of Mathematical Psychology*, 20 (1), 1–15.

Jedidi, Kamel and Wayne S. DeSarbo (1991), "A Stochastic Multidimensional Scaling Procedure for the Spatial Representation of Three-Mode, Three-Way Pick Any/J Data," *Psychometrika*, 56 (3), 471–94.

Kass, Robert E. and Adrian E. Raftery (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90 (430), 773–95.

Krumhansl, Carol D. (1978), "Concerning the Applicability of Geometric Models to Similarity Data: The Interrelationship Between Similarity and Spatial Density," *Psychological Review*, 85 (5), 445–63.

Kruskal, Joseph B. and Myron Wish (1983), *Multidimensional Scaling*. Newbury Park, CA: Sage Publications.

Linden, Greg D., Brent R. Smith, and Nida K. Zada (2005), "Use of Product Viewing Histories of Users to Identify Related Products," U.S. Patent Number: 6,912,505 B2, filed March 29, 2001, and issued June 28, 2005.

Lynch, John G., Jr., and Dan Ariely (2000), "Wine Online: Search Costs Affect Competition on Price, Quality, and Distribution," *Marketing Science*, 19 (1), 83–103.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2003), "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science*, 22 (1), 58–84.

Moe, Wendy (2006), "An Empirical Two-Stage Choice Model with Varying Decision Rules Applied to Internet Clickstream Data," *Journal of Marketing Research*, 43 (November), 680–92.

Newman, Joseph W. and Bradley D. Lockeman (1975), "Measuring Prepurchase Information Seeking," *Journal of Consumer Research*, 2 (3), 216–22.

——— and Richard Staelin (1972), "Prepurchase Information Seeking for New Cars and Major Household Appliances," *Journal of Marketing Research*, 9 (August), 249–57.

Okada, Akinori and Tadashi Imaizumi (1987), "Nonmetric Multidimensional Scaling of Asymmetric Proximities," *Behaviormetrika*, 14 (21), 87–96.

Payne, John W. (1976), "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance*, 16 (August), 366–87.

Roberts, John H. and James M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28 (November), 429–40.

Saburi, S. and N. Chino (2008), "A Maximum Likelihood Method for an Asymmetric MDS Model," *Computational Statistics and Data Analysis*, 52 (10), 4673–84.

Siddarth, S., Randolph E. Bucklin, and Donald G. Morrison (1995), "Making the Cut: Modeling and Analyzing Choice Set Restriction in Scanner Panel Data," *Journal of Marketing Research*, 32 (August), 255–66.

Shugan, Steven M. (1980), "The Cost of Thinking," *Journal of Consumer Research*, 7 (September), 99–111.

Stigler, George (1961), "The Economics of Information," *Journal of Political Economy*, 69 (3), 213–25.

Swait, Joffre and Tülin Erdem (2007), "Brand Effects on Choice and Choice Set Formation Under Uncertainty," *Marketing Science*, 26 (5), 666–78.

Urban, Glen L., John S. Hulland, and Bruce D. Weinberg (1993), "Premarket Forecasting for New Consumer Durable Goods: Modeling Categorization, Elimination, and Consideration Phenomena," *Journal of Marketing*, 57 (April), 47–63.

Wedel, Michel, and Wayne S. DeSarbo (1996), "An Exponential-Family Multidimensional Scaling Mixture Methodology," *Journal of Business & Economic Statistics*, 14 (4), 447–59.

Yee, Michael, Ely Dahan, John R. Hauser, and James Orlin (2007), "Greedoid-Based Noncompensatory Inference," *Marketing Science*, 26 (4), 532–49.